

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号
特表2003-523188
(P2003-523188A)

(43) 公表日 平成15年8月5日 (2003.8.5)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード (参考)
C 1 2 N 15/09	Z N A	A 6 1 K 39/21	4 B 0 2 4
A 6 1 K 39/21		A 6 1 P 31/14	4 B 0 6 4
A 6 1 P 31/14		31/18	4 B 0 6 5
31/18		C 0 7 K 14/155	4 C 0 8 5
C 0 7 K 14/155		16/08	4 H 0 4 5
審査請求 未請求 予備審査請求 有 (全 109 頁) 最終頁に続く			

(21) 出願番号 特願2001-560222(P2001-560222)
(86) (22) 出願日 平成13年2月16日 (2001.2.16)
(85) 翻訳文提出日 平成14年8月19日 (2002.8.19)
(86) 国際出願番号 PCT/US 01/05288
(87) 国際公開番号 WO 01/060838
(87) 国際公開日 平成13年8月23日 (2001.8.23)
(31) 優先権主張番号 60/183,659
(32) 優先日 平成12年2月18日 (2000.2.18)
(33) 優先権主張国 米国 (US)

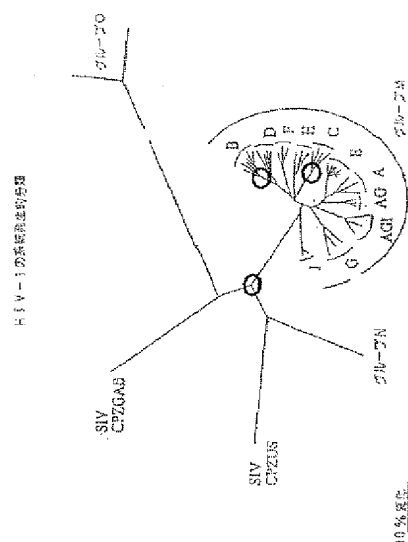
(71) 出願人 ユニバーシティ・オブ・ワシントン
アメリカ合衆国ワシントン州98105, シア
トル, ノース・イースト・フォーティフィ
フス・ストリート 1107
(72) 発明者 ミュリンズ, ジェイムズ アイ.
アメリカ合衆国, ワシントン 98105, シ
アトル, ノースイースト, イースト ロー
レルハースト ドライブ 3134
(72) 発明者 ロドリゴ, アレン ジー.
ニュージーランド国, オークランド, マレ
イズ ベイ, シートン ロード 8
(74) 代理人 弁理士 石田 敬 (外4名)

最終頁に続く

(54) 【発明の名称】 AIDS先祖ウイルスおよびワクチン

(57) 【要約】

本発明は、先祖HIV核酸およびアミノ酸配列、このよう
な配列の製造方法およびその使用、例えば予防的および
診断的使用に向けられる。



【特許請求の範囲】

【請求項1】 高度多様性ウイルス株、サブタイプまたはグループの決定された始祖配列である、単離された先祖ウイルス遺伝子配列およびその断片。

【請求項2】 先祖ウイルス遺伝子配列がHIV-1先祖ウイルス遺伝子配列、HIV-2先祖ウイルス遺伝子配列またはC型肝炎先祖ウイルス遺伝子配列である請求項1記載の配列。

【請求項3】 先祖ウイルス遺伝子配列がHIV-1サブタイプA、B、C、D、E、F、G、H、J、AGまたはAGI；HIV-1グループM、NまたはO；あるいはHIV-2サブタイプAまたはBのものである請求項1記載の配列。

【請求項4】 先祖ウイルス遺伝子配列が広範分散性HIV-1変異体、地理的制限性HIV-1変異体、広範分散性HIV-2変異体または地理的制限性HIV-2変異体のものである請求項1記載の配列。

【請求項5】 先祖ウイルス遺伝子配列がenv遺伝子またはgag遺伝子である請求項1記載の配列。

【請求項6】 先祖ウイルス遺伝子配列が、平均して、任意の他の変異体に対してより任意の所与の循環性ウイルスの遺伝子配列に対してより密接に関連する請求項1記載の配列。

【請求項7】 配列番号1、配列番号3、配列番号5または配列番号6で記述される配列と少なくとも70%の同一性を有し、そして任意の循環性変異体と100%の同一性を有しない請求項1記載の配列。

【請求項8】 配列番号2または配列番号4の先祖タンパク質をコードする請求項1記載の配列。

【請求項9】 HIV-1、HIV-2またはC型肝炎からの単離された先祖タンパク質またはその断片。

【請求項10】 配列番号2または配列番号4の連続配列を含む請求項9記載の単離された先祖タンパク質。

【請求項11】 HIV-1サブタイプA、B、C、D、E、F、G、H、J、AGまたはAGI；グループM、NまたはO；あるいはHIV-2サブタイプAまたはBの先祖タンパク質である請求項9記載の単離された先祖タンパク質。

【請求項12】 HIV-1サブタイプBenv先祖タンパク質またはHIV-1サブタイプCenv先祖タンパク質の少なくとも10連続アミノ酸である請求項11記載の単離された先祖タンパク質。

【請求項13】 gagまたはenvタンパク質である請求項9記載の単離された先祖タンパク質。

【請求項14】 以下の作用可能に連結された要素：
転写プロモーター；
先祖タンパク質をコードする核酸；および
転写ターミネーター
を含む単離された発現構築物。

【請求項15】 核酸が配列番号2または配列番号4をコードする請求項14記載の発現構築物。

【請求項16】 核酸が配列番号1、配列番号3、配列番号5または配列番号6として記述される配列である請求項14記載の発現構築物。

【請求項17】 核酸配列が宿主細胞中での発現のために最適化される請求項13記載の発現構築物。

【請求項18】 前記転写プロモーターが異種プロモーターである請求項14記載の発現構築物。

【請求項19】 前記プロモーターがサイトメガロウイルスプロモーターである請求項18記載の発現構築物。

【請求項20】 請求項14の発現構築物で形質転換またはトランスフェクトされる培養された原核または真核生物細胞。

【請求項21】 哺乳類細胞である請求項20記載の真核生物細胞。

【請求項22】 核酸が配列番号2または配列番号4の先祖タンパク質をコードする請求項20記載の真核生物細胞。

【請求項23】 核酸が配列番号1、配列番号3、配列番号5または配列番号6として記述される配列である請求項20記載の真核生物細胞。

【請求項24】 大腸菌細胞である請求項20記載の原核生物細胞。

【請求項25】 ビール酵母菌*S. cerevisiae*細胞である請求項20記載の真

核生物細胞。

【請求項26】 ヒト細胞である請求項20記載の真核生物細胞。

【請求項27】 請求項14の発現構築物を含むベクター。

【請求項28】 核酸配列がセムリキ森林熱 (Semliki Forest Virus) ウィルスレプリコンと作用可能に連結され、その結果生じた組換えレプリコンがサイトメガロウイルスプロモーターと作用可能に連結される請求項27記載のベクター。

【請求項29】 請求項27のベクターを含む単離された宿主細胞。

【請求項30】 哺乳類における免疫応答を誘導するための組成物であって、高度多様性ウイルス先祖タンパク質または先祖タンパク質の免疫原性断片を含む組成物。

【請求項31】 前記断片が配列番号2または配列番号4で記述される配列に由来する請求項30記載の組成物。

【請求項32】 前記ウイルス先祖タンパク質がHIV-1またはHIV-2からである請求項30記載の組成物。

【請求項33】 ワクチンである請求項30記載の組成物。

【請求項34】 ウイルス先祖タンパク質と特異的に結合し、そして複数の循環性子孫ウイルス先祖タンパク質と特異的に結合する単離された抗体。

【請求項35】 先祖タンパク質がHIV-1、HIV-2またはC型肝炎からである請求項34記載の抗体。

【請求項36】 モノクローナル抗体またはその抗原結合断片である請求項34記載の抗体。

【請求項37】 ヒト化モノクローナル抗体である請求項34記載の単離された抗体。

【請求項38】 抗体またはその抗原結合断片が一本鎖抗体、キメラ抗体、一本重鎖抗体、抗原結合F(ab')₂断片、抗原結合Fab'断片、抗原結合Fab断片または抗原結合Fv断片である請求項34記載の抗体。

【請求項39】 先祖ウイルスアミノ酸配列の製造方法であって、
(a) 高度多様性ウイルスの循環性ウイルス配列を選択し、

(b) 循環性ウイルス配列の進化系統樹の進化中心を代表する先祖ウイルス配列である循環性ウイルス配列の最新の共通の祖先である先祖ウイルス配列を最大見込み系統発生分析により確定し、そして

(c) 循環性ウイルス配列のいずれかと100%同一ではないが、しかしその推定アミノ酸配列がそれらのうちのいずれかと少なくとも70%同一であるウイルス配列を合成する、

ことを包含する方法。

【請求項40】 循環性ウイルス配列がHIV-感染患者からである請求項39記載の方法。

【請求項41】 循環性ウイルス配列がHIV-1、HIV-2またはC型肝炎からである請求項39記載の方法。

【請求項42】 免疫原性に関して一検定で断片を試験することをさらに包含する請求項39記載の方法。

【請求項43】 最大見込み系統発生分析が合一見込み分析を包含する請求項39記載の方法。

【発明の詳細な説明】

【0001】

発明の背景

HIV-1は、ワクチン開発のための非常に難しい標的であることが立証されている。HIV-1感染に対する防御免疫の免疫相関物は、依然として不明である。ウイルスは感染個体中で持続的に複製して、活発な体液性および細胞性免疫応答の発生にもかかわらず、容赦なく疾患を引き起こす。HIV-1は感染中に急速に突然変異を起こして、免疫認識を逃れ得るウイルスの生成を引き起こす。他の高度多様性ウイルス（例えばインフルエンザ）と違って、一原型株が連続均一系統により置き換えられる一連の変異体であるとは思えない。むしろ、多数のHIV-1感染個体からサンプリングされたウイルス配列の進化系統樹は、系統樹の中心からほぼ等距離の変異体のほとんどに関してスターバーストパターンを形成する。HIV-1ウイルスはまた、後に個体中で複製し得る潜在プロウイルスDNAとして無期限に存続し得る。

【0002】

今日、いくつかのHIV-1ワクチンアプローチが開発中であり、各々がそれ自体の相対的強さおよび弱さを有する。これらのアプローチは、生弱毒ワクチン、アジュバントペプチドを用いた不活性化ウイルスおよびサブユニットワクチン、生ベクターベースのワクチンならびにDNAワクチンの開発を包含する。エンベロープ糖タンパク質は、それらが理想的免疫原ではないことが明白になるまで、それらの表面曝露によるワクチンレジメンにおける主要抗原と考えられていた。これは、これらのウイルスの進化を推進する免疫学的選択力の予測結果である。天然感染における不十分な免疫原性を指図するエンベロープ糖タンパク質の同一特徴がワクチン開発を妨げる、と考えられる。しかしながらワクチン製法の修正は、これらの問題を克服し得る。例えば融合競合免疫原を有する感染個体からの一次単離物の好結果の中和（マウスにおける）についての近年の報告は、この考えを支持する。

【0003】

別のアプローチは、ワクチン製法においてHIV-1の天然単離物を使用すること

である。しかしながら初期変異体の同定は、AIDS流行を開始しそうな保存検体からでさえ、非常に見込みがない。天然単離物は、ワクチン候補のための理想である特徴（例えばエピトープ）を具体化することも見込みがない。さらに任意の所定天然ウイルス単離物は、その特定のヒト宿主内の特異的相互作用による適応を反映する特徴を有する。これらの個体特異的特徴は、ウイルスの全てのまたはほとんどの系統に見出されるとは予測されず、したがって個々の単離物をベースにしたワクチンは広範囲の循環性ウイルスに対して有効であるとは考えられない。

【0004】

別のアプローチは、HIV-1試験に対する広範な防御を引き出すための努力に際してワクチン製法においてできるだけ多数の多様性HIV-1単離物を包含することである。先ず、1つまたはそれ以上の系統が、HIVの多数の循環系統の中から選択される。このアプローチの利点は、このような系統が感染形態の成育可能ウイルスであることが既知であるということである。しかしながらこのような系統は循環中の他の系統に対して遺伝的に全く似ておらず、したがって広範な防御を引き出せない。関連アプローチは、循環系統をまたはデータベース中の系統を基礎にしたコンセンサス配列を構築することである。コンセンサス配列は循環系統から遺伝的意味でそれほど離れていないと考えられるが、しかし任意の実際のウイルスの概算はなく、したがって広範な防御を提供し得ない。

したがってHIV感染を予防および治療するためのワクチン開発のための候補配列を同定する新規の有効な方法に対する当業界での必要性が存在する。本発明は、このそしてその他の必要を満たす。

【0005】

発明の要約

本発明は、先祖ウイルス遺伝子配列およびウイルス先祖タンパク質配列を確定するための組成物および方法を提供する。一つの観点では、高度多様性ウイルス、例えばHIV-1、HIV-2またはC型肝炎に関する先祖ウイルス配列を確定するために用いられ得る計算方法が提供される。これらの計算方法は、最大見込み系統発生分析により先祖ウイルス配列を確定するために循環性ウイルスの試料を用いる。

【0006】

先祖ウイルス配列は、例えばHIV-1先祖ウイルス遺伝子配列、HIV-2先祖ウイルス遺伝子配列またはC型肝炎先祖ウイルス遺伝子配列であり得る。別の実施態様では、先祖ウイルス遺伝子配列は、HIV-1サブタイプA、B、C、D、E、F、G、H、J、AGまたはAGI；HIV-1群M、NまたはO；あるいはHIV-2サブタイプAまたはBのものである。先祖ウイルス遺伝子配列は、広範分散性HIV-1変異体、地理的制限性HIV-1変異体、広範分散性HIV-2変異体または地理的制限性HIV-2変異体のものでもあり得る。典型的には、先祖遺伝子はenv遺伝子またはgag遺伝子である。

【0007】

先祖ウイルス遺伝子配列は、平均して任意の他の変異体より任意の所与の循環性ウイルスの遺伝子配列とより密接に関連する。いくつかの実施態様では、先祖ウイルス遺伝子は、配列番号1、配列番号3、配列番号5または配列番号6で記述される配列と少なくとも70%の同一性を有するが、しかし任意の循環変異体と100%の同一性を有するというわけではない。

【0008】

一つの観点において、本発明は、HIV-1サブタイプBのenv遺伝子に関する先祖配列を提供する。HIV-1サブタイプBは、西半球および欧州において大半の感染を生じる。確定先祖ウイルス配列は、平均して、任意の他の変異体より任意の所与の循環性ウイルスとより密接に関連する。Env先祖遺伝子配列は、884アミノ酸長である、envの遺伝子産物gp160に関するオープンリーディングフレームをコードする。

【0009】

別の観点では、本発明は、HIV-1サブタイプCのenv遺伝子に関する先祖配列を提供する。サブタイプCは、世界中で最も流布しているサブタイプである。この配列は、平均して、任意の他の変異体より任意の所与の循環性ウイルスとより密接に関連する。この配列は、853アミノ酸長である、envの遺伝子産物gp160に関するオープンリーディングフレームをコードする。

【0010】

単離されたHIV先祖タンパク質またはその断片も提供される。単離された先祖

タンパク質は、例えばHIV-1サブタイプBenv先祖タンパク質（配列番号2）またはHIV-1サブタイプCenv先祖タンパク質（配列番号4）の連続配列であり得る。先祖タンパク質は、HIV-1サブタイプA、B、C、D、E、F、G、H、J、AGまたはAGI；群M、NまたはO；あるいはHIV-2サブタイプAまたはBのものでもあり得る。

【0011】

本発明は、他の先祖ウイルス配列を確定するための計算方法も提供する。計算方法は、例えば他のHIVサブタイプ、例えば発展途上国において広範に蔓延しているHIV-1サブタイプEに関する先祖ウイルス配列を確定することにおよび得る。計算方法は、すべての既知のならびに新規出現高度多様性ウイルス、例えばHIV-1系統、サブタイプおよびグループに関する先祖ウイルス配列を確定することにも及ぶ。例えば先祖ウイルス配列は、タイまたはブラジルにおけるHIV-1-B、中国、インド、南アフリカまたはブラジルにおけるHIV-1-C等に関して確定され得る。他の実施態様では、先祖ウイルス配列は、HIV-1nef遺伝子またはポリペプチド、pol遺伝子またはポリペプチドあるいはその他の補助遺伝子またはポリペプチドに関して確定される。

【0012】

本発明は、発現構築物、例えば転写プロモーター；先祖タンパク質をコードする核酸；および転写ターミネーターも提供する。核酸は、例えばHIV-1先祖タンパク質（例えば配列番号2または配列番号4）をコードし得る。核酸は、例えばHIV-1サブタイプBまたはCenv遺伝子配列（例えば配列番号1、配列番号3、配列番号5または配列番号6）であり得る。一実施態様では、核酸配列は宿主細胞中での発現のために最適化される。プロモーターは異種プロモーター、例えばサイトメガロウイルスプロモーターであり得る。発現構築物は、原核または真核生物細胞中で発現され得る。適切な細胞としては、例えば哺乳類細胞、人細胞、大腸菌細胞およびビール酵母菌*S. cerevisiae*細胞が挙げられる。一実施態様では、発現構築物は、セムリキ森林熱ウイルスレプリコンと作用可能に連結された核酸配列を有し、この場合、その結果生じた組換えレプリコンはサイトメガロウイルスプロモーターと作用可能に連結される。

【0013】

別の観点では、組成物は哺乳類における免疫応答を誘導するために提供され、組成物はウイルス先祖タンパク質または先祖タンパク質の免疫原性断片を含有する。先祖タンパク質は、HIV-1サブタイプBまたはCenv先祖タンパク質から、あるいはその他のHIV-1、HIV-2またはC型肝炎先祖タンパク質から得られる。組成物は、高度多様性1型ヒト免疫不全ウイルスワクチン (HIV-1) により感染に対して防御するための例えばAIDSワクチン、あるいはHIV-2またはC型肝炎感染を防御するためのワクチンとして用いられ得る。先祖ウイルス配列は、HIV-1グループ先祖 (例えばグループM)、HIV-1サブタイプ (例えばB、CまたはE)、広範蔓延性変異体 (widely spread variant)、地理的制限性変異体 (geographically-restricted variant) または新規発現変異体であり得る。

【0014】

別の観点では、ウイルス先祖タンパク質と特異的に結合し、そして複数の循環子孫ウイルス先祖タンパク質と特異的に結合する単離抗体が提供される。先祖タンパク質は、例えばHIV-1、HIV-2またはC型肝炎からであり得る。抗体は、モノクローナル抗体またはその抗原結合断片であり得る。一実施態様では、抗体はヒト化モノクローナル抗体である。その他の適切な抗体またはその抗原結合断片は、一本鎖抗体、一本重鎖抗体、抗原結合F(ab')₂断片、抗原結合Fab'断片、抗原結合Fab断片または抗原結合Fv断片であり得る。

【0015】

先祖ウイルス配列を確定することのほかに、本発明は、先祖ウイルス配列を基礎にした免疫原性組成物を調製し、試験するための方法も提供する。特定の実施態様では、免疫原性組成物 (先祖ウイルス配列に基づいた) が調製され、適切なモデル、例えばマウスモデルまたはサルーヒト免疫不全ウイルス (SHIV) アカゲザルモデルを用いて、哺乳類に投与される。免疫原性組成物は、単離先祖ウイルス遺伝子配列またはポリペプチド配列あるいはその一部を用いて調製され得る。

さらに別の観点では、先祖ウイルス配列を基礎にした核酸、ペプチドまたは抗体を用いて、被験者におけるHIVおよび/またはAIDSを検出するための診断方法が提供される。

【0016】

特定の実施態様の説明

本発明をさらに詳細に記述する前に、本明細書中で以後用いられるようなある種の用語の定義を記述することは、そのさらなる理解に有用であると思われる。

【0017】

定義

別記しない限り、本明細書中で用いられるすべての技術および科学用語は、本発明が属する業界の当業者により一般的に理解されたのと同じ意味を有する。本明細書中に記載されたものと同様の任意の方法および材料が本発明の実施または試験に用いられ得るが、しかし例示的方法および材料のみを記載する。本発明の目的のために、以下の用語は下記のように定義する。

【0018】

本発明の状況において、「先祖配列」とは、典型的には、平均して任意の他の変異体より任意の所定変異体とより密接に関連するものである決定された始祖配列 (determined founder sequence) を指す。「先祖ウイルス配列」とは、典型的には、平均して任意の他の変異体に対してより任意の所定の循環性ウイルスに対してより密接に関連するものである決定された始祖配列を指す。「先祖ウイルス配列」は、循環性ウイルスの核酸および／またはアミノ酸配列を用いて最大見込み系統発生分析（本明細書中により十分に記載されているような）の適用により確定される。「先祖ウイルス」とは、「先祖ウイルス配列」を含むウイルスである。「先祖タンパク質」とは、アミノ酸先祖ウイルス配列を有するタンパク質、ポリペプチドまたはペプチドである。

【0019】

「循環性ウイルス」という用語は、感染個体中に見出されるウイルスを指す。

「変異体」という用語は、1つまたはそれ以上のヌクレオチドまたはアミノ酸による他のウイルス、遺伝子または遺伝子産物と配列が異なるウイルス、遺伝子または遺伝子産物を指す。

【0020】

「免疫学的な」または「免疫応答」という用語は、レシピエント被験者におけるHIVペプチドに対して向けられる有益な体液性（即ち抗体媒介性）および／ま

たは細胞性（即ち抗原特異的T細胞またはそれらの分泌産物により媒介される）応答の発現を指す。このような応答は、特に、免疫原の投与により誘導される活性応答であり得る。細胞免疫応答は、抗原特異的CD4⁺Tヘルパー細胞（即ちヘルパーTリンパ球）および／またはCD8⁺細胞傷害性T細胞を活性化するために、クラスIまたはクラスII MHC分子に関連したエピトープの提示により引き出される。

【0021】

細胞媒介性免疫学的応答の存在は、例えばCD4⁺T細胞の増殖検定（即ち、HTL（ヘルパーTリンパ球）応答を測定する）により、あるいはCTL（細胞傷害性Tリンパ球）検定により確定され得る（例えば、Burke et al., J. Inf. Dis. 170:1110-19(1994); Tigges et al., J. Immunol. 156:3901-10(1996)参照）。免疫原の防御的または治療的作用に対する体液性および細胞性応答の相対的寄与は、免疫感作同質遺伝子的動物（immunized syngeneic animal）からIgGおよびT細胞を別々に単離し、二次被験者における防御的または治療的作用を測定することにより区別され得る。例えばエフェクター細胞は削除され、その結果生じる応答が分析され得る（例えば、Schmitz et al., Science 283:857-60(1999); Jin et al., J Exp. Med. 189:991-98(1999)参照）。

【0022】

「抗体」とは、分析物（抗原）を特異的に結合し、認識する単数または複数の免疫グロブリン遺伝子またはその断片により実質的にコードされるポリペプチドを指す。認識免疫グロブリン遺伝子としては、 κ 、 λ 、 α 、 γ 、 δ 、 ϵ および μ 定常部遺伝子、ならびに無数の免疫グロブリン可変部遺伝子が挙げられる。軽鎖は、 κ または λ として分類される。重鎖は、 γ 、 μ 、 α 、 δ または ϵ として分類され、これらは順次、それぞれ免疫グロブリンクラスIgG、IgM、IgA、IgDおよびIgEを限定する。

【0023】

例示的免疫グロブリン（抗体）の構造単位は、四量体を含む。各四量体は、ポリペプチド鎖の2つの同一対で構成され、各対は1つの「軽（約25 kD）」鎖および1つの「重」鎖（約50~70 kD）を有する。各鎖のN末端は、主に抗原認識に関与する約100~110またはそれ以上のアミノ酸の可変部を有する。可変軽鎖（VL）

および可変重鎖 (VH) という用語は、それぞれこれらの軽および重鎖を指す。

【0024】

抗体は、例えば無傷免疫グロブリンとして、または種々のペプチダーゼを用いた消化により産生される多数の十分特性化された抗原結合断片として存在する。例えばペプシンは、ヒンジ部のジスルフィド結合より下で抗体を消化して、それ自体はジスルフィド結合により VH-CH1 に連結された軽鎖である Fab の二量体である $F(ab')_2$ 断片を産生する。 $F(ab')_2$ 断片は、穏やかな条件下で還元されて、ヒンジ部のジスルフィド結合を壊し、それにより $F(ab')_2$ 二量体を Fab' 単量体に転換し得る。

【0025】

Fab' 単量体は、本質的にはヒンジ部の一部を伴う Fab である (Fundamental Immunology, Third Edition, W.E.Paul (ed.), Raven Press, N.Y. (1993) 参照)。種々の抗体断片は無傷抗体の消化により限定されるが、一方、このような断片は化学的にまたは組換え DNA 法を利用することにより de novo で合成され得る、と当業者は理解する。したがって抗体という用語は、本明細書中で用いる場合、抗体断片、例えば一本鎖抗体、抗原結合 $F(ab')_2$ 断片、抗原結合 Fab' 断片、抗原結合 Fab 断片、抗原結合 Fv 断片、一本重鎖またはキメラ抗体も含む。このような抗体は、全抗体の修飾により産生され得るか、あるいは組換え DNA 法を用いて de novo で合成され得る。

【0026】

「生物学的試料」という用語は、ゲノムまたはウイルス DNA あるいはその他の核酸 (例えば mRNA、ウイルス RNA 等) もしくはタンパク質を有する任意の組織または液体試料を指す。「生物学的試料」はさらに、無細胞ウイルスを含有する流体、例えば血清および血漿を含み、そして正常健常細胞および HIV 感染が疑われる細胞もともに含む。

【0027】

「核酸」という用語は、一本鎖または二本鎖形態のデオキシリボヌクレオチドまたはリボヌクレオチドおよびそのポリマーを指す。特定の限定されない限り、当該用語は、参照核酸と同様の結合特性を有する天然ヌクレオチドの既知の類

似体を含む核酸を包含する。別記しない限り、特定の核酸配列は、その保存的修飾化変異体（例えば縮重コドン置換）および相補配列、ならびに明瞭に示された配列も暗黙的に包含する。特に縮重コドン置換は、1つまたはそれ以上の選定（またはすべての）コドンの第三位置が、混合塩基および／またはデオキシイノシン残基で置換される配列を生成することにより達成され得る（例えば、Batzner et al., *Nucleic Acid Res.* 19:5081(1991); Ohtsuka et al., *J. Biol. Chem.* 260:2605-08(1985); Rossolini et al., *Mol. Cell. Probes* 8:91-98(1994) 参照）。

【0028】

核酸は、少なくとも10連続ヌクレオチドの断片（例えばハイブリッド可能部分）も包含する；別の実施態様では、核酸は少なくとも25ヌクレオチド、50ヌクレオチド、100ヌクレオチド、150ヌクレオチド、200ヌクレオチドあるいは250ヌクレオチドまでまたはそれ以上を含む。「核酸」という用語は、遺伝子、cDNAおよび遺伝子によりコードされるmRNAと互換的に用いられる。

【0029】

本明細書中で用いる場合、「核酸プローブ」とは、1つまたはそれ以上の種類の化学結合により、通常は相補塩基対合により、例えば水素結合形成により、相補配列の標的核酸（例えばHIV-1核酸）と結合し得る核酸と定義される。本明細書中で用いる場合、プローブは天然（例えばA、G、CまたはT）または修飾化塩基（例えば7-デアザグアノシン、イノシン等）を包含し得る。さらにプローブ中の塩基は、それがハイブリダイゼーションを妨げない限り、ホスホジエステル結合以外の結合により連結され得る。したがって、例えばプローブは、組成塩基がホスホジエステル結合よりむしろペプチド結合により連結されるペプチド核酸であり得る。ハイブリダイゼーション条件の緊縮度によるレベルで、プローブ配列との完全相補性を欠く標的配列をプローブが結合し得る、ということは、当業者により理解される。

【0030】

核酸プローブは、DNAまたはRNA断片であり得る。DNA断片は、例えばプラスミドDNAを消化することにより、PCRの使用により、または化学合成により、例えば

BeaucageとCarruthers (Tetrahedron Lett. 22:1859-62(1981))により記載されているホスホロアミダイト法により、あるいはMatteucci等 (J. Am. Chem. Soc. 103:3185(1981)) によるトリエステル法により調製され得る。次に、所望により、適切な条件下で化学合成一本鎖と一緒にアニーリングすることにより、または適切なプライマー配列を有するDNAポリメラーゼを用いて相補鎖を合成することにより、二本鎖断片が得られる。核酸プローブに関する特定の配列が示される場合、相補鎖も同定され、包含されると理解される。相補鎖は、標的が二本鎖化核酸である状況で等しく良好に機能する。

【0031】

「標識化核酸プローブ」とは、プローブに結合された標識の存在を検出することによりプローブの存在が検出されるように、共有的に、リンカーにより、またはイオン、ファンデルワールスまたは水素結合により、標識と結合される核酸プローブである。

「作用可能に連結される」という用語は、核酸発現制御配列（例えばプロモーター、シグナル配列、または転写因子結合部位の任意のアレイ）と二次核酸配列との間の機能的連結を指すが、この場合、発現制御配列は、二次配列に対応する核酸の転写および／または翻訳に影響を及ぼす。

【0032】

「増幅プライマー」は、選定核酸配列の増幅のための基準として役立ち得る天然または類似ヌクレオチドを含む核酸、典型的にはオリゴヌクレオチドである。それらの例としては、例えばポリメラーゼ連鎖反応プライマーおよびリガーゼ連鎖反応オリゴヌクレオチドの両方が挙げられる。

「ポリペプチド」、「ペプチド」および「タンパク質」という用語は、アミノ酸残基のポリマーを指すために本明細書中で互換的に用いられる。本用語は、1つまたはそれ以上のアミノ酸残基が対応する天然アミノ酸ならびに天然アミノ酸ポリマーおよび非天然アミノ酸ポリマーの人工化学模倣物であるアミノ酸ポリマーに当てはまる。

【0033】

「アミノ酸」または「アミノ酸残基」という用語は、本明細書中で用いられる

場合、さらに以下に記載されるような天然L-アミノ酸またはD-アミノ酸を指す。アミノ酸に関して一般的に用いられる一文字および三文字略号が本明細書中で用いられる（例えば、Alberts et al., *Molecular Biology of the Cell*, Garland Publishing, Inc., New York(3d ed. 1994); Creighton, *Proteins*, W.H. Freeman and Company(1984)参照）。

【0034】

「保存的置換」とは、タンパク質を説明する場合、タンパク質の活性を実質的に変えるとは思われないタンパク質のアミノ酸組成の変化を指す。したがって、特定アミノ酸配列の「保存的修飾化変異」とは、重要でさえあるアミノ酸の置換が活性を実質的に変えないような、タンパク質活性に関して重要であるとは思われないアミノ酸のアミノ酸置換、あるいは同様の特性（例えば、酸性、塩基性、正または負荷電、極性または非極性等）を有する他のアミノ酸を用いたアミノ酸の置換を指す。しばしば機能的に同様であるアミノ酸を提示する保存的置換表は、当業界で周知である（例えば、Creighton, *Proteins*, W.H. Freeman and Company(1984)参照）。さらに、単一アミノ酸またはコード配列中の小バーセンテージのアミノ酸を変更、付加または欠失する個々の置換、欠失または付加も「保存的修飾化変異」である。

【0035】

2つまたはそれ以上の核酸またはポリペプチド配列の状況において、「同一の」または「同一性パーセント」という用語は、以下の配列比較アルゴリズムの1つを用いて、または手動アラインメントおよび視覚的検査により測定した場合、比較ウィンドウまたは意図された領域に亘って最大対応性に関して比較し、整列した場合に、同一であるか、または同一であるアミノ酸残基またはヌクレオチドの特定パーセンテージ（即ち特定領域に亘って60%同一性、任意に65%、70%、75%、80%、85%、90%または95%同一性）を有する2つまたはそれ以上の配列または亜配列を指す。このような配列はその場合、「実質的に同一である」といわれる。この定義は、試験配列の相補体も指す。任意に、同一性は、少なくとも約30アミノ酸またはヌクレオチド長である領域に亘って、典型的には50、75または150アミノ酸またはヌクレオチドである領域に亘って存在する。一実施態様では、配列

はコード領域の全長に亘って実質的に同一である。

【0036】

2つまたはそれ以上のポリペプチド配列の状況において、「類似性」または「類似性パーセント」という用語は、以下の配列比較アルゴリズムの1つを用いて、または手動アラインメントおよび視覚的検査により測定した場合、比較ウィンドウまたは意図された領域に亘って最大対応性に関して比較し、整列した場合に、前記の保存的アミノ酸置換で限定されるように同一であるかまたは類似する特定パーセンテージのアミノ酸残基（即ち特定領域に亘って少なくとも60%、任意に65%、70%、75%、80%、85%、90%または95%類似する）を有する2つまたはそれ以上の配列または重配列を指す。このような配列はその場合、「実質的に類似する」といわれる。任意に、この同一性は、少なくとも約25アミノ酸長である領域に亘って、さらに好ましくは約50、75または100アミノ酸長である領域に亘って存在する。

【0037】

配列比較に関しては、典型的には一配列は、試験配列が比較される参照配列として作用する。配列比較アルゴリズムを用いた場合、試験および参照配列は、典型的にはコンピューターに入力され、必要ならば重配列同等物が指示され、そして配列アルゴリズムプログラムパラメーターが指示される。配列比較アルゴリズムは次に、指定プログラムパラメーターに基づいて、参照配列と比較して試験配列（単数または複数）に関する配列同一性パーセントを算定する。

【0038】

比較のための配列の最適アラインメントは、例えばSmithとWaterman (Adv. Appl. Math. 2:482(1981)) の局所相同アルゴリズムにより、NeedlemanとWunsch (J. Mol. Biol. 48:443(1970)) の相同アラインメントアルゴリズムにより、PearsonとLipman (Pro. Natl. Acad. Sci. USA 85:2444(1988)) の同一性に関する検索法により、これらのアルゴリズムのコンピューター処理実行 (Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr. Madison, WIのGAP、BESTFIT、FASTAおよびTFASTA)により、または視覚的検査（一般に、A usubel et al., Current protocols in Molecular Biology, John Wiley and So

ns, New York(1996)参照) により実行され得る。

【0039】

有用なアルゴリズムの一例は、PILEUPである。PILEUPは、関係および配列同一性パーセントを示すために累進対合式アラインメントを用いて関連配列の一群から多重配列アラインメントを作製する。それは、アラインメントを作製するために用いられる群生関係を示す系統樹または系統発生樹もプロットする。PILEUPは、FengとDoolittle (J. Mol. Evol. 35:351-60(1987)) の累進アラインメント法の単純化を用いる。用いられる方法は、HigginsとSharp (Gene 73: 237-44 (1988); CABIOS 5: 151-53(1989)) により記載されたCLUSTAL法に類似する。プログラムは300配列まで並べ得るし、各々、5,000ヌクレオチドまたはアミノ酸の最大長を有する。多重アラインメント手法は、2つの最もよく似た配列の対合方式を用いて開始し、2つの整列化配列の一群を生じる。

【0040】

この一群はその後、次のもっとも関連した配列または整列化配列団に整列される。2つの群の配列は、2つの個々の配列の対合式アラインメントの単純延長により整列される。最終アラインメントは、一連の累進対合式アラインメントにより達成される。プログラムは、配列比較の領域に関して特定の配列およびそれらのアミノ酸またはヌクレオチド同等物を指示することにより、そしてプログラムパラメーターを指示することにより、実行される。例えば参照配列は、以下のパラメーター：デフォルトgap重量 (3.00)、デフォルトgap長重量 (0.10) および計量末端gapを用いて、配列同一性パーセント関係を確定するために、他の試験配列と比較され得る。

【0041】

配列同一性パーセントおよび配列類似性を確定するのに適したアルゴリズムの別の例は、BLASTアルゴリズムであり、これはAltschul等 (J. Mol. Biol. 215:403-10(1990)) に記載されている。BLAST分析を実施するためのソフトウェアは、National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>) を通じて公的に利用可能である。このアルゴリズムは、データベース中の同一長のワードを用いて整列される場合、何らかの正評価閾値得点Tに適合する

かまたはそれを満たす疑問配列中の長さ W の短いワードを同定することにより、高得点配列対 (HSP) を先ず同定することを包含する。 T は、隣接ワード得点閾値として示される (Altschul et al., 上記)。

【0042】

これらの初期隣接ワードヒットは、それらを含むより長いHSPを見出すための検索を開始するための核として作用する。ワードヒットは次に、累積アライメント得点が増大され得る限り、各配列に沿って両方向に延長される。累積得点は、ヌクレオチド配列に関して、パラメーター M (マッチング残基対に関する応報得点; 常に >0) および N (ミスマッチング残基に関するペナルティー得点; 常に <0) を用いて算定される。アミノ酸配列に関しては、採点マトリックスを用いて、累積得点を算定する。各方向のワードヒットの延長は、累積アライメント得点とその最大達成値から量 X を減じる場合; 1つまたはそれ以上の負の採点残基アライメントの蓄積のために、累積得点がゼロまたはそれ以下になる場合; あるいは配列の末端が到達される場合に、停止される。

【0043】

BLASTアルゴリズムパラメーター W 、 T および X は、アライメントの感度および速度を決定する。BLASTNプログラム (ヌクレオチド配列に関して) は、デフォルトとして、3のワード長 (W)、10の予測数量 (E)、100の遮断、 $M=5$ 、 $N=-4$ および両鎖の比較を用いる。アミノ酸配列に関しては、BLASTNプログラムは、デフォルトとして、3のワード長 (W)、10の予測数量 (E)、およびBLOSUM62採点マトリックスを用いる (Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915(1989)参照)。

【0044】

配列同一性パーセントを算定するほかに、BLASTアルゴリズムは、2つの配列間の類似性の統計学的分析も実施する (例えば、Karlin and Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-87(1993)参照)。BLASTアルゴリズムにより提供される類似性の一測定値は最小合計確率 ($P(N)$) であり、これは、2つのヌクレオチドまたはアミノ酸配列間の適正が偶然により起こる確率の指標を提供する。例えば核酸は、参照核酸との試験核酸の比較における最小合計確率が典型的には約0.

35～約0.1である場合には、参照配列に類似すると考えられる。

【0045】

2つの核酸が実質的に同一であるという別の指標は、2つの分子が緊縮条件下で互いにハイブリダイズするということである。「～と特異的にハイブリダイズする」という語句は、その配列が複合混合物（例えば全細胞性）DNAまたはRNA中に存在する場合に緊縮条件下での特定のヌクレオチド配列のみとの分子の結合、重複またはハイブリダイゼーションを指す。「実質的に結合する（単数または複数）」とは、プローブ核酸と標的核酸との間の相補的ハイブリダイゼーションを指し、標的ポリヌクレオチド配列の所望の検出を達成するためにハイブリダイゼーション培地の緊縮を低減することにより収容され得る小不適正を包含する。

【0046】

サザンおよびノーザンハイブリダイゼーションのような核酸ハイブリダイゼーション実験の状況における「緊縮ハイブリダイゼーション条件」および「緊縮ハイブリダイゼーション洗浄条件」は配列依存性であり、異なる環境パラメーター下で異なる。より長い配列は、高温で特異的にハイブリダイズする。核酸のハイブリダイゼーションに対する広範な指針は、Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, part I, chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays," Elsevier, N.Y.(1993)に見出される。一般に、高緊縮ハイブリダイゼーションおよび洗浄条件は、限定イオン強度およびpHでの特定配列に関する融点 (T_m) より約5℃低いよう選定される。典型的には、「緊縮条件」下では、プローブはその標的配列とハイブリダイズするが、しかし他の配列とはハイブリダイズしない。

【0047】

T_m は、標的配列の50%が完全適正プローブとハイブリダイズする温度（限定イオン強度およびpH下）である。極緊縮条件は、特定プローブに関する T_m と等しいよう選択される。サザンまたはノーザンブロットのフィルター上に100より多い相補残基を有する相補核酸のハイブリダイゼーションに関する緊縮ハイブリダイゼーション条件の一例は、4～6× SSCまたはSSPEを含有する水性溶液中で42℃

または65~68℃で4~6x SSCまたはSSPE中の50%ホルムアミドである。高緊縮洗浄条件の一例は、72℃で約15分間の0.15 MNaClである。緊縮洗浄条件の一例は、65℃で15分間の0.2x SSC洗浄である（一般的に、Sambrook et al., Molecular cloning, A Laboratory Manual, 2nd ed., Cold Spring Harbor Publish., Cold Spring Harbor, NY (1989)参照）。

【0048】

しばしば、高緊縮洗浄は、バックグラウンドプローブシグナルを除去するために低緊縮に先行される。例えば100より多いヌクレオチドの二重鎖に関する中緊縮洗浄の一例は、45℃で15分間の1x SSCである。例えば100より多いヌクレオチドの二重鎖に関する低緊縮洗浄の一例は、40℃で15分間の4~6x SSCである。短いプローブ（例えば約10~50ヌクレオチド）に関しては、緊縮条件は典型的には、pH7.0~8.3で約1.0 M未満のNaイオンの塩濃度、典型的には約0.01~1.0 MNaイオン濃度（またはその他の塩）を包含し、温度は典型的には少なくとも約30℃である。

【0049】

緊縮条件は、脱安定剤、例えばホルムアミドの付加によっても達成され得る。概して、特定ハイブリダイゼーション検定における非関連プローブに関して観察されたものの2倍（またはそれ以上）のシグナル対ノイズ比は、特異的ハイブリダイゼーションの検出を示す。緊縮条件下で互いにハイブリダイズしない核酸は、それらがコードするポリペプチドが実質的に同一である場合には、実質的に同一である。これは、例えば遺伝暗号により可能にされる最大コドン縮重を用いて核酸のコピーが作製される場合に起こる。

【0050】

2つの核酸またはポリペプチドが実質的に同一であるというさらなる指標は、一次核酸によりコードされるポリペプチドが、二次核酸によりコードされるポリペプチドに対して生じた抗体と免疫学的に交差はのせいであるかまたはそれと特異的に結合することである。したがって、ポリペプチドは典型的には、例えば2つのペプチドが保存的置換によってのみ異なる二次ポリペプチドと実質的に同一である。

【0051】

「抗体と特異的に（または選択的に）結合する」または「～と特異的に（または選択的に）免疫反応性である」という語句は、タンパク質またはペプチドに言及する場合、タンパク質およびその他の生物学的物質の異種集団の存在下でのタンパク質の存在の決定因である結合反応を指す。したがって指定イムノアッセイ条件かでは、特異化抗体は特定タンパク質と結合し、試料中に存在する他のタンパク質と有意量で結合しない。このような条件下でのタンパク質との特異的結合は、特定タンパク質に対するその特異性のために選定される抗体を要する。例えば本発明の核酸のいずれかによりコードされるアミノ酸配列を有するタンパク質に対して生じる抗体は、そのタンパク質とは特異的に免疫反応性であるが、多形変異体を除いて他のタンパク質とは特異的免疫反応性でない抗体を得るよう選定され得る。

【0052】

種々のイムノアッセイフォーマットを用いて、特定タンパク質と特異的に免疫反応性である抗体を選定し得る。例えば固相ELISAイムノアッセイ、ウエスタンブロットまたは免疫組織化学法は、タンパク質と特異的に免疫反応性であるモノクローナル抗体を選定するためにルーチンに用いられる（例えば特異的免疫反応性を確定するために用いられ得るイムノアッセイフォーマットおよび条件の説明に関しては、Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, N.Y.(1988)参照）。典型的には、特異的または選択的反応は、バックグラウンドシグナルまたはノイズの少なくとも2倍、さらに典型的にはバックグラウンドの10~100倍である。

【0053】

「免疫原性組成物」という用語は、特異的免疫原に対して抗体を生じる免疫応答または細胞媒介性免疫応答を引き出す組成物を指す。免疫原性組成物は、注射用として、液体溶液、懸濁液、乳濁液等として調製され得る。

「ワクチン」という用語は、疾患、特にウイルス疾患に対する防御を付与するための、霊長類、特にヒト宿主であり得る宿主へのin vivo投与のための免疫原性組成物を指す。

【0054】

「単離された」という用語は、その天然細胞環境から取り出されたウイルス、核酸またはポリペプチドを指す。単離ウイルス、核酸またはポリペプチドは、典型的には、細胞核酸、ポリペプチドおよびその他の構成成分から少なくとも部分的に精製される。

本発明の状況では、「合一事象」とは、それらの最新の共通先祖の時点での系統学上の2つの直系の結合を指す。

「合一間隔」とは、合一事象間の時間を説明する。各合一間隔に関する予測時間は指数関数的に分布され、 $n \ll N$ に関して平均 $E[t_{n, n-1}] = 2N/n(n-1)$ 生成である。

【0055】

先祖配列の系統発生的確定

一つの観点において、先祖配列を確定するための計算方法が提供される。このような方法は、例えばウイルスに関する先祖配列を確定するために用いられ得る。これらの計算方法は、典型的には高度多様性ウイルス集団として存在するウイルスの先祖配列を確定するために用いられる。例えばいくつかの高度多様性ウイルス（例えばHIV-1、HIV-2、C型肝炎等）は、一原型系統が連続均一系統により置き換えられる連続変異体により進化するとは考えられない。代わりに、ウイルス配列の進化系統樹は「スターバーストパターン」を形成し、ほとんどの変異体がスターバーストの中心からほぼ等距離にある。このスターバーストパターンは、多重多様性循環系統が共通先祖から進化することを示す。このような高度多様性ウイルス、例えばHIV-1、HIV-2、C型肝炎およびその他のウイルスに関する先祖配列を確定するために、計算方法が用いられ得る。

【0056】

先祖配列の確定方法は、典型的には循環性ウイルスの核酸配列を基礎にする。ウイルス核酸配列は複製されるので、それは複製過程におけるエラーによる塩基変化を獲得する。例えばいくつかの核酸配列が複製される場合、チミン (T) はその正常相補体であるシトシン (C) よりむしろグアニン (G) と結合し得る。これらの塩基変化（または突然変異）のほとんどはその後の複製事象で再生されな

いが、しかし一定の割合の突然変異は子孫配列に受け渡される。多くの複製サイクルを用いるほど、核酸配列が獲得する突然変異は多い。1つまたはそれ以上の突然変異を保有する核酸配列が2つの別個の直系を生じる場合には、その結果生じる2つの直系は同一親核酸配列を共有し、同一親突然変異（単数または複数）を有する。これらの直系の「歴史」を遡及すると、それらは、2つの直系が共通先祖から生じた共通分枝点を有する。同様に、現在の循環性ウイルス核酸配列の歴史を遡ると、この歴史の分枝点も、単一先祖が子孫直系を生じた点（節点として指示されている）に対応する。

【0057】

本発明の計算方法は、最大見込みの原理を基礎にしており、循環性ウイルスの核酸配列の試料を用いる。試料中のウイルスの配列は、典型的には共通の特徴を共有し、例えば同一ウイルス系統、サブタイプまたは群からである。系統発生は、ウイルス核酸の複製に際してヌクレオチド置換の確率を特定する進化のモデルを用いることにより構築される。ヌクレオチドが異なる配列中の位置（即ち突然変異の部位）で、観察されたウイルス配列を得る確率が最大になるよう、当該方法は、節点（即ち直系の分枝点）にヌクレオチドの1つを割り当てる。

【0058】

節点へのヌクレオチドの割当は、単数または複数の予測系統発生を基礎にする。各データ組に関しては、異なるウイルス系統、サブタイプまたは群からのいくつかの配列は、当該配列を捜し出すための外群として用いられる。配列置換のモデル、次いで最大見込み系統発生は、各データ組（例えばサブタイプおよび外群）に関して確定される。最大見込み系統発生は、試料中に観察核酸配列を生じる最高確率を有する。最大見込み系統発生は、基礎節点での配列は、先祖配列（または最新の共通先祖）と呼ばれる（例えば、図1および2参照）。したがってこの先祖配列は、試料内の異なる配列からほぼ等距離にある。

【0059】

最大見込み系統発生は、循環性ウイルスの配列の試料を用いる。循環性ウイルスの配列は、例えば感染者の血液、組織またはその他の生物学的試料から核酸を抽出し、そしてウイルス核酸をシーケンシングすることにより確定され得る（例

えば、Sambrook et al., Molecular cloning, A Laboratory Manual, 2nd ed., Cold Spring Harbor Publish., Cold Spring Harbor, N.Y.(1989); Kriegler, Gene Transfer and Expression: A Laboratory Manual, W.H. Freeman, N.Y.(1990); Ausubel et al., 上記、参照)。

【0060】

一実施態様では、抽出ウイルス核酸はポリメラーゼ連鎖反応により増幅され、次にDNAがシーケンシングされる。循環性ウイルスの試料は、保存生物学的試料からおよび／または循環性ウイルスの試料から予測的に（例えば、インド対エチオピアにおけるHIV-1サブタイプCをサンプリングする）得られる。ウイルス配列は、データベース（例えばGenBankおよびLos Alamos配列データベース）からも同定され得る。

【0061】

一旦循環性ウイルスの試料が収集されれば（典型的には約20～約50試料）、本発明の計算方法を用いて、1つまたはそれ以上の遺伝子に関する核酸配列が分析される。一方法では、配列中の任意の指定部位に関して、系統樹上の全ての節点にヌクレオチドが割り当てられる。循環性ウイルスの観察配列を得る確率を最大にするすべての節点に関するヌクレオチドの形状が確定される。この方法を用いて、全節点にまたがる状態の連結見込みが最大にされる。

【0062】

第二の方法は、系統樹上の指定ヌクレオチド部位および指定節点に関して、循環性ウイルスの観察配列を得る確率を最大にし、系統樹上の他の節点でのヌクレオチドのすべての考え得る割当を可能にするヌクレオチドを選択することである。この第二の方法は、特定の割当の限界見込みを最大にする。これらの方法に関しては、先祖配列の再構築（即ち先祖状態）は、しかしながら単一確定配列のみを生じる必要がない。それらの見込み順にランクされた多数の先祖配列を選択することができる。

【0063】

HIV集団に関しては、モデリングの第二層が最大見込み系統発生分析に付加され、特に当該層は分析に用いられる進化のモデルに付加される。この第二層は、

合一見込み分析を基礎にする。合一とは、配列の系統の数学的説明であり、集団に作用する過程を考慮に入れる。これらの過程が何らかの確実性を伴って既知である場合、合一の使用は、各種類の系統樹に従来の確率を割り当てるために用いられ得る。系統樹の見込みを用いて考慮すると、データが示されれば確定系統発生樹が正しいという後期確率が確定され得る。一旦系統樹が選定されれば、先祖状態は、前記と同様に確定され得る。したがって合一見込み分析は、先祖ウイルス配列（例えば、始祖または最新の共通先祖（MRCA）の配列）の配列を確定するためにも適用され得る。

【0064】

典型的実施態様では、最大見込み系統発生分析は、先祖配列（例えば先祖ウイルス配列）を決定するために適用される。典型的には、ウイルス系統、サブタイプまたは群のような共通の特徴を有する20~50核酸配列試料が用いられる（例えば、同一サブタイプの世界的多様性を包含する試料）。他のウイルス（例えば別の系統、サブタイプまたは群）からの付加的配列が得られ、分析中のウイルス配列を搜し出すための外群として用いられる。ウイルス配列の試料は、データベース（例えばGenBankおよびLos Alamos配列データベース）から、あるいは配列情報の同様の供給源から同定された目下の循環性ウイルスから確定される。

【0065】

配列は、CLUSTALW (Thompson et al., Nucleic Acids Res. 22:4673-80(1994)) (この記載内容は、参照により本明細書中に含まれる) を用いて整列され、これらのアラインメントは、GDE (Smith et al., CABIOS 10:671-75(1994)) (この記載内容は、参照により本明細書中に含まれる) を用いて精製される。アミノ酸配列は、核酸配列から翻訳される。ギャップは、それらがコドン間に挿入されるよう操作される。このアラインメント（アラインメントI）は、明瞭に整列され得ない領域が除去されるよう、系統発生的分析のために修飾され (Learn et al., J. Virol. 70:5720-30(1996)) (この記載内容は、参照により本明細書中に含まれる)、アラインメントIIを生じる。

【0066】

これらの配列（アラインメントII）に関する系統発生および先祖状態再構築の

ための適切な進化モデルは、Modeltest 3.0 (Posada and Crandall, Bioinformatics 14:817-8(1998))(この記載内容は、参照により本明細書中に含まれる)で実行される場合、Akaike情報判定基準 (AIC)(Akaike, IEEE Trans. Autom. Contr. 19: 716-23 (1974))(この記載内容は、参照により本明細書中に含まれる)を用いて選定される。

【0067】

例えばサブタイプC先祖配列のための分析に関しては、最適モデルは、転写の両クラスに関しては等速度であり、トランスバージョンの4つのクラス全てに関しては異なる速度であって、不変部位、ならびに可変部位の部位-部位速度変動性のX分布を伴う (TVM+I+Gモデルとして言及される)。この場合のモデルのパラメーターは、例えば平衡ヌクレオチド頻度： $f_A=0.3576$ 、 $f_C=0.1829$ 、 $f_G=0.2314$ 、 $f_T=0.2290$ ；不変部位の割合=0.2447；X分布の形状パラメーター (α) =0.7623；速度マトリックス (R) マトリックス値： $R_{A \rightarrow C}=1.7502$ 、 $R_{A \rightarrow G}=R_{C \rightarrow T}=4.1332$ 、 $R_{A \rightarrow T}=0.6825$ 、 $R_{C \rightarrow G}=0.6549$ 、 $R_{G \rightarrow T}=1$ 。

【0068】

配列 (アラインメントII) に関する進化系統樹は、PAUPバージョン4.0b (Swofford, PAUP 4.0: Phylogenetic Analysis Using Parsimony (And Other Methods); Sinauer Associates, Inc. (2000))(この記載内容は、参照により本明細書中に含まれる)で実行される場合、最大見込み概算 (MLE) 法を用いて推定される。例えば、HIV-1サブタイプC配列に関しては、10の異なる亜系統樹-剪定-再移植 (SPR) 発見的検索が、各々異なる無作為付加順を用いて実施され得る。先祖ウイルスヌクレオチド配列は、系統発生を用いて基礎節点での配列であることが確定され、データベース (アラインメントII)、ならびに前記のTVM+I+Gモデルからの配列は、限界見込み概算 (下記参照) を用いて確定される。

【0069】

いくつかの場合、確定配列は可変部 (例えば、HIV-1-Cに関する可変部V1、V2、V4およびV5) の一部に関する先祖配列を含み得ないし、あるいはいくつかの短い領域は明瞭には整列され得ない。以下の手法は、最適には、高可変部 (例えばアラインメントIから欠失されたもの) を含めた完全配列に関するアミノ酸配列

を予測するために用いられ得る。確定先祖配列は視覚的にアラインメントIと一列に整列されて、GDE (Smith et al., 上記) を用いて翻訳される。高可変部は完全コドンとして欠失され得るため、翻訳リーディングフレームは保存され、コドンは維持され得る。アラインメントIIから欠失された領域に関する先祖アミノ酸配列は視覚的に予測され、コンピュータプログラムMacCladeバージョン3.08a (Maddison and Maddison. MacClade-Analysis of Phylogeny and Character Evolution -Version 3. Sinauer Associates, Inc.(1992)) を用いたこれらの部位に関する最節約ベースの配列再構築を用いて精製される。

【0070】

先祖アミノ酸配列は、任意に、特定細胞型における発現のために最適化される。アミノ酸配列は、例えばWisconsin配列分析パッケージ (GCG)、バージョン10のBACKTRANSLATEプログラムならびにコドン使用データベース ([http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Homo+sapiens+\[gbpri\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Homo+sapiens+[gbpri])) (これらの記載内容はともに、参照により本明細書中に含まれる) からのヒト遺伝子コドン表を用いて、ある種の細胞型 (例えばヒト細胞) 中での発現のために最適化されたDNA配列に転換され得る。

【0071】

最適化配列は、非最適化先祖配列と同一のアミノ酸配列を、当該遺伝子 (例えばenv遺伝子) に関してコードする。最適化配列を有する合成ウイルスは、RNA二次構造的特徴 (例えばHIV-1のRev応答素子 (RRE)) 等の存在下での異なるリーディングフレーム中の補助遺伝子の崩壊のために完全に機能的であるというわけではない。最適化過程は補助遺伝子 (例えばHIV-1のvpu、tatおよびrev遺伝子) のコード領域に影響を及ぼし得るし、RNA二次構造を破壊し得る。したがって先祖配列は、半最適化され得る。半最適化配列は、他の特徴に亘らない配列の部分に関する最適化配列を有するが、この場合、非最適化先祖配列が代わりに用いられる。例えばHIV-1先祖配列に関しては、最適化先祖配列は、vpu、tat、revおよびRRE領域に亘らない配列の部分に関して用いられ、一方、「非最適化」先祖配列は、vpu、tat、revおよびRRE領域を重複する配列の部分に関して用いられる。

【0072】

HIV先祖ウイルス配列の系統発生的決定

先祖ウイルス配列は、1型HIV (HIV-1)、2型HIV (HIV-2) またはその他のHIVウイルス、例えばHIV-1サブタイプ、HIV-2サブタイプに関する、その他のHIVサブタイプに関する、出現性HIVサブタイプに関する、ならびにHIV変異体、例えば広範分散性または地理的単離変異体からの任意の単数または複数の遺伝子に関して確定され得る。

【0073】

例えば先祖ウイルス遺伝子配列は、HIV-1、例えばHIV-1サブタイプA、B、C、D、E、F、G、H、J、AG、AGIのenvおよびgag遺伝子に関して、ならびに群M、N、Oに関して、あるいはHIV-2ウイルスまたはHIV-2サブタイプAまたはBに関して確定され得る。特定の実施態様では、先祖ウイルス配列は、HIV-1サブタイプBおよび／またはCのenv遺伝子に関して、またはサブタイプBおよび／またはCからのgag遺伝子に関して確定される。他の実施態様では、先祖ウイルス配列は、その他のHIV遺伝子またはポリペプチド、例えばnef、polまたはその他の補助遺伝子またはポリペプチドに関して確定される。

【0074】

現在および／または以前の循環性ウイルスからの選定HIV-1またはHIV-2遺伝子の核酸配列は、既存のデータベース（例えばGenBankまたはLos Alamos配列データベース）から同定され得る。循環性ウイルスの配列は、組換えDNA法によっても確定され得る（例えば、Sambrook et al., Molecular cloning, A Laboratory Manual, 2nd ed., Cold Spring Harbor Publish., Cold Spring Harbor, N.Y.(1989); Kriegler, Gene Transfer and Expression: A Laboratory Manual, W.H. Freeman, N.Y.(1990); Ausubel et al., 上記、参照）。

【0075】

各データ組に関しては、異なるウイルス系統、サブタイプまたは群からのいくつかの配列は、当該配列を捜し出すための外群として用いられる。配列置換のモデル、次いで最大見込み系統発生は、各データ組（例えばサブタイプおよび外群）に関して確定される。先祖ウイルス配列は、変異体配列の基礎節点での配列として確定される（例えば、図1および2参照）。したがってこの先祖配列は、サブ

タイプ内の異なる配列からほぼ等距離にある。

【0076】

一実施態様では、先祖HIV-1群M、サブタイプB、env配列は、41の異なる単離物を用いて確定された（確定核酸およびアミノ酸配列は、それぞれ表1および2（配列番号1および配列番号2）に図示されている）。図2を参照しながら、38のサブタイプB配列および3つのサブタイプD（外群）配列を用いて、サブタイプB配列を搜し出した。サブタイプB配列は、9カ国からのものであり、サブタイプB多様性の広範な試料を示す：オーストラリア、8配列；中国、1配列；フランス、5配列；ガボン、1配列；ドイツ、2配列；英国、2配列；オランダ、2配列；スペイン、1配列；米国、15配列。

【0077】

確定先祖タンパク質は、884アミノ酸長である。この先祖ウイルス配列とそれを確定するために用いられた循環系統との間の距離は、平均で12.3%（8.0～21.0%の範囲）であったが、一方、利用可能な検体は、互いに17.3%異なっていた（13.3～23.2%の範囲）。したがって先祖配列は、平均で、任意の他の変異体より任意の所定循環性ウイルスにより密接に関連する。他のサブタイプB系統と比較した場合、先祖配列は、アミノ酸レベルで94.6%の同一性を有して、USAD8 (Theodore et al., AIDS Res. Human Retrovir. 12: 191-94 (1996)) に最も類似する。

【0078】

意外なことに、HIV-1サブタイプBenv遺伝子の確定先祖ウイルス配列は、抗原提示のために処理された場合、広範囲の種々の免疫学的に活性ペプチドをコードする。ほぼすべての既知のサブタイプB CTLエピトープコンセンサスアミノ酸（387/390; 99.23%）が、サブタイプB、gp160配列に関して確定先祖ウイルス配列中表示される。これに対比して、HIV-1サブタイプBの大半のその他の変異体は、95%より低いエピトープ配列保存を示す（これは先祖ウイルス配列の必要な特徴というわけではないが、しかしHIV-1の急速延長の結果である）。したがって、このサブタイプB先祖タンパク質に対する免疫原性組成物は、同一サブタイプPのHIV-1単離物に対して広範な中和抗体を引き出す。このサブタイプB先祖タンパク質に対する免疫原性組成物は、抗原特異的T細胞により媒介される広範な細胞応

答も引き出す。

【0079】

別の実施態様では、同様の計算方法を用いて、HIV-1サブタイプCenv遺伝子配列の先祖ウイルス配列を確定した。HIV-1サブタイプCは、発展途上国において蔓延している。サブタイプCは、世界的に最も一般的なサブタイプであって、HIV-1感染のおよそ30%に関与し、アフリカ、インドおよび中国における流行の主要構成成分である。HIV-1群MサブタイプCenv遺伝子に関する先祖ウイルス配列は、57の異なる単離物（39のサブタイプC配列および18の外群配列（2つは他の群Mサブタイプの各々から）；図8）を用いて確定された。確定アミノ酸配列は、表4（配列番号4）に示されている。ヒト細胞中での発現のために最適化された確定核酸配列は、表3（配列番号3）に示されている。

【0080】

サブタイプC配列は、12のアフリカおよびアジア諸国からであり、世界規模のサブタイプC多様性の広範な試料を表示する：ボツワナ、8配列；ブラジル、2配列；ブルンジ、8配列；中華人民共和国、1配列；ジブチ、2配列；エチオピア、1配列；インド、8配列；マラウイ、3配列；セネガル、1配列；ソマリア、1配列；ウガンダ、1配列；およびザンビア、3配列。確定先祖タンパク質は、853アミノ酸長である。この先祖ウイルス配列とそれを確定するために用いられる循環系統との間の距離は、平均で11.7%（9.3～14.3%の範囲）であり、一方、利用可能な検体は、平均で互いに16.6%異なっていた（7.1～21.7%の範囲）。したがって先祖タンパク質配列は、平均して、任意の他の変異体より任意の指定循環性ウイルスにより密接に関連する。他のサブタイプC系統と比較した場合、先祖配列はMW965 (Gao et al., J Virol. 70:1651-67(1996)) と最も類似し、アミノ酸レベルで89.5%の同一性を有する。

【0081】

意外なことに、確定先祖ウイルス配列は、抗原提示のために処理された場合、広範な種々の免疫学的活性ペプチドをコードする。ほぼすべての既知のサブタイプC CTLエピトープコンセンサス配列（389/396；98.23%）が、サブタイプCgp160配列に関する確定先祖ウイルス配列中に表示される。これに対比して、HIV-1サ

ブタイプCの典型的変異体（先祖配列を確定するために用いられるもの）は、95.19%未満のエピトープ配列保存（平均90.36%、64.56~95.19%の範囲）を有する。したがって、このサブタイプC先祖ウイルス配列に対するワクチンは同一サブタイプのHIV-1単離物に対する広範な中和抗体を引き出す。このサブタイプC先祖タンパク質に対する免疫原性組成物も、抗原特異的T細胞により媒介される広範な細胞性応答を引き出す。

【0082】

HIV先祖配列に関する最適化および半最適化配列も提供される。先祖ウイルス配列は、特定の宿主細胞中での配列のために最適化され得る。最適化先祖配列は非最適化配列と同一の遺伝子に関するアミノ酸配列をコードし、一方、最適化配列は、異なるリーディングフレーム、RNA二次構造の崩壊等における補助遺伝子の崩壊のために、合成ウイルス中で十分には機能し得ない。例えばHIV-1env配列の最適化は、vpu、tatおよび／またはrev、および／またはRNA二次構造Rev応答性素子（RRE）に関する補助遺伝子を崩壊し得る。

【0083】

半最適化配列は、他の遺伝子、RNA二次構造等に及ばない配列の一部に関する最適化配列を用いることにより調製される。このような特徴を重複する配列の一部に関して、「非最適化」先祖配列が用いられる（例えばvpu、tat、revおよび／またはRREを重複する領域に関して）。特定の実施態様では、HIV-1サブタイプBおよびCに関する半最適化先祖ウイルス配列が提供される（表5（配列番号5）および表6（配列番号6）参照）。

【0084】

他の実施態様では、先祖ウイルス配列は、広範循環変異体または地理的制限性変異体に関して確定される。例えば、広範囲に分散される（例えば、多数の国にまたは明白な地理的境界のない地域に存在する）HIV-1サブタイプについて試料が収集され得る。同様に、地理的に（例えば、国、地域またはその他の物理的限定領域に）制限されたHIV-1サブタイプの試料が収集される。試料中の遺伝子（例えばgagまたはenv）の配列は、組換えDNA法（例えば）により、あるいはデータベース中の情報から確定される。典型的には、試料の数は、それらの一般的利

用可能性、ならびにウイルスが当該領域中で循環されていた時間（例えばウイルスが循環されていた時間が長いほど、多様性は大きく、試料から収集される情報は多い）によって、約20～約50の範囲である。次に核酸またはアミノ酸である先祖ウイルス配列が、本明細書中に記載した計算方法を用いて確定される。

【0085】

先祖ウイルス配列をコードする核酸

本明細書に記載された方法により先祖ウイルス配列が確定されれば、組換えDNA法を用いて、当該先祖ウイルス配列をコードする核酸を調製し得る。適切な方法としては、(1) 先祖ウイルス配列に最も類似する既存ウイルス系統を修飾すること；(2) より短いオリゴヌクレオチド（例えば160～200ヌクレオチド長）を連結することにより先祖ウイルス配列をコードする核酸を合成すること；または(3) これらの方法の組合せ（例えば、より発散性配列を*de novo*で合成しながら、先祖ウイルス配列との極高度類似性を有する断片を用いて既存配列を修飾することによる）が挙げられるが、これらに限定されない。

【0086】

核酸配列が生成され、ルーチン技法を用いて操作され得る（例えばSambrook et al., 上記、Kriegler, 上記、Ausubel et al., 上記、参照）。別記しない限り、酵素はすべて、メーカーの使用説明書にしたがって用いられる。

【0087】

典型的実施態様では、先祖ウイルス配列をコードする核酸は、長いオリゴヌクレオチドを連結することにより合成される。核酸を*de novo*で合成することにより、所望の特徴が遺伝子中に容易に組入れられる。このような特徴としては、核酸配列のさらなる操作を可能にするのに便利な制限部位、ポリペプチド配列の免疫原性を付与し得る*in vivo*発現レベルを大いに強化するためのコドン頻度（例えばヒトコドン頻度）の最適化等が挙げられるが、これらに限定されない。長いオリゴヌクレオチドは、固相法を用いて極低誤差率で合成され得る。5'末端の両方で20～25ヌクレオチド相補配列を用いて設計された長いオリゴヌクレオチドは、DNAポリメラーゼ、DNAリガーゼなどを用いて連結され得る。必要ならば、合成核酸の配列は、DNA配列分析により立証され得る。

【0088】

市販されていないオリゴヌクレオチドは、化学的に合成され得る。適切な方法としては、例えばBeaucageとCarruthers (Tetrahedron Lett. 22:1859-62(1981))により最初に記載された固相ホスホロアミダイトトリエステル法、ならびに自動合成機の使用（例えば、Needham Van Devanter et al., Nucleic Acids Res. 12:6159-68(1984)参照）が挙げられる。オリゴヌクレオチドの精製は、例えばネイティブアクリルアミドゲル電気泳動によるか、またはPearsonとReanier (J. Chrom. 255:137-49(1983))に記載されたような陰イオン交換HPLCによる。

【0089】

核酸の配列は、例えばMaxam等 (Methods in Enzymology 65: 499-560(1980))の化学分解法または二本鎖鋳型をシーケンシングするためのチェインターミネーター法（例えば、Wallace et al., Gene 16: 21-26(1981)参照）を用いて立証され得る。サザンブロットハイブリダイゼーション技法は、Southern等 (J. Mol. Biol. 98: 503(1975))、Sambrook等（上記）またはAusubel等（上記）にしたがって実行され得る。

【0090】

先祖ウイルス配列の発現

先祖ウイルス配列をコードする核酸は、適切な発現ベクター（即ち、挿入ポリペプチドコード配列の転写および翻訳に必要な素子を含有するベクター）中に挿入され得る。種々の宿主ベクター系は、ポリペプチドコード配列（単数または複数）を発現するために利用され得る。これらの例としては、例えばウイルス（例えばワクシニアウイルス、アデノウイルス、シンドビスウイルス、ベネズエラウマ脳脊髄炎（VEE）ウイルス等）を感染させた哺乳類細胞系、ウイルス（例えばバキュロウイルス）を感染させた昆虫細胞系、微生物、例えば酵母ベクターを含有する酵母、またはバクテリオファージDNA、プラスミドDNAまたはコスミドDNAで形質転換された細菌が挙げられる。

【0091】

ベクターの発現素子は、それらの強度および特異性が変化する。利用される宿主ベクター系によって、多数の適切な転写および翻訳素子のいずれかが用いら

れ得る。特定の実施態様では、先祖ウイルス配列は、ヒト細胞、その他の哺乳類細胞、酵母または細菌中で発現される。さらに別の実施態様では、配列の免疫学的に活性な領域を含む先祖ウイルス配列の断片が発現される。

【0092】

発現ベクター中葉の先祖ウイルス配列をコードする核酸の挿入のために、任意の適切な方法が用いられ得る。適切な発現ベクターは、典型的には、適切な転写および翻訳制御シグナルを含む。適切な方法としては、*in vitro*組換えDNAおよび合成技法、ならびに*in vivo*組換え技法（遺伝子組換え）が挙げられる。核酸配列の発現は、コード化核酸が組換えDNA分子で形質転換された宿主中で発現されるように、二次核酸配列により調節され得る。例えば、先祖ウイルス配列の発現は、当業界で既知の適切なプロモーター／エンハンサー素子により制御され得る。

【0093】

適切なプロモーターとしては、例えばSV40初期プロモーター領域 (Beoist and Chambon, Nature 290:304-10(1981))、ラウス肉腫ウイルスの3'長末端反復中に含入されるプロモーター (Yamamoto et al., Cell 22:787-97(1980))、ヘルペスチミジンキナーゼプロモーター (Wagner et al., Proc. Natl. Acad. Sci. U.S.A. 78: 1441-45 (1981))、サイトメガロウイルスプロモーター、翻訳延長因子EF-1 α プロモーター、メタロチオネイン遺伝子の調節配列 (Brinster et al., Nature 296:39-42(1982))、原核生物プロモーター、例えば β -ラクタマーゼプロモーター (Villa-Komaroff et al., Proc. Natl. Acad. Sci. USA 75:3727-31(1978)) またはtacプロモーター (deBoer et al., Proc. Natl. Acad. Sci. USA 80: 21-25(1983))、植物発現ベクター、例えばカリフラワーモザイクウイルス35SRNAプロモーター (Gardner et al., Nucl. Acids Res. 9:2871-88(1981))、ならびに光合成酵素リブローズビスホスフェートカルボキシラーゼのプロモーター (Herrera-Estrella et al., Nature 310:115-20(1984))、酵母またはその他の真菌からのプロモーター素子、例えばGAL7およびGAL4プロモーター、ADH（アルコールデヒドロゲナーゼ）プロモーター、PGK（ホスホグリセロールキナーゼ）プロモーター、アルカリ性ホスファターゼプロモーター等が挙げられる。

【0094】

その他の哺乳類プロモーターの例としては、例えば、組織特異性を示す以下の動物転写制御領域が挙げられる：膵臓腺房細胞中で活性であるエラスターゼI遺伝子制御領域 (Swift et al., Cell 38:639-46 (1984); Ornitz et al., Cold Spring Harbor Symp. Quant. Biol. 50:399-409(1986); MacDonald, Hepatology 7(1 Suppl.): 42S-51S(1987)) ; 膵臓 β 細胞中で活性であるインスリン遺伝子制御領域 (Hanahan, Nature 315:115-22(1985)) ; リンパ細胞中で活性である免疫グロブリン遺伝子制御領域 (Grosschedl et al., Cell 38: 647-58(1984); Adams et al., Nature 318:533-38(1985); Alexander et al., Mol. Cell. Biol. 7: 1436-44(1987)) ; 精巣、乳房、リンパおよびマスト細胞中で活性であるマウス乳癌ウイルス制御領域 (Leder et al., Cell 45:485-95(1986)) ;

【0095】

肝臓中で活性であるアルブミン遺伝子制御領域 (Pinkert et al., Genes Dev. 1:268-76(1987)) ; 肝臓中で活性である α -フェトプロテイン遺伝子制御領域 (Krumlauf et al., Mol. Cell. Biol. 5:1639-48(1985); Hammer et al., Science 235:53-58(1987)) ; 肝臓中で活性である α 1-抗トリプシン遺伝子制御領域 (Kelsey et al., Genes and Devel. 1:161-71(1987)) ; 骨髄細胞中で活性である β -グロブリン遺伝子制御領域 (Magram et al., Nature 315:338-40(1985); Kollias et al., Cell 46:89-94(1986)) ; 脳の乏突起膠細胞中で活性であるミエリン塩基性タンパク質遺伝子制御領域 (Readhead et al., Cell 48:703-12(1987)) ; 骨格筋中で活性であるミオシン軽鎖-2遺伝子制御領域 (Shani, Nature 314: 283-86(1985)) ; ならびに視床下部中で活性であるゴナドトロピン放出ホルモン遺伝子制御領域 (Mason et al., Science 234:1372-78(1986)) 。

【0096】

特定の実施態様では、先祖ウイルス配列コード核酸と操作可能的に連結されるプロモーター、1つまたはそれ以上の複製起点、ならびに任意に1つまたはそれ以上の選択可能マーカー（例えば抗生物質耐性遺伝子）を含むベクターが用いられる。選択可能マーカーとしては、例えばアンピシリン、テトラサイクリン、ネオマイシン、G418等に対する耐性を付与するものが挙げられる。発現構築物は、例

例えば先祖ウイルス配列をコードする核酸をpRSECT発現ベクターの制限部位中にサブクローニングすることにより作製され得る。このような構築物は、発現ポリペプチドのアフィニティー精製のためのヒスチジンアミノ末端フラッグ配列を用いたT7プロモーターの制御下での先祖ウイルス配列の発現を可能にする。

【0097】

例示的实施態様では、以下でさらに考察されるように、最大タンパク質発現を達成するために極高効率RNA／タンパク質発現構成成分（例えばセムリキ森林熱ウイルスから）を有する高効率性DNA伝達ベクター（pJW4304 SV40/EBVベクター）を用いる高効率性発現系が用いられ得る。pJW4304SV40/EBVは、Robinson等（Ann. New York Acad. Sci. 27:209-11(1995)）およびYasutomi等（J. Virol. 70: 678-81(1996)）により記載されているpJW4303から調製される。

【0098】

先祖ウイルス配列を発現する発現ベクター／宿主系は、当業者に周知の一般的アプローチにより、例えば（a）核酸ハイブリダイゼーション、（b）「マーカー」遺伝子機能の存在または非存在、（c）挿入配列の発現、あるいは（d）標準組換えDNA法による形質転換化細胞のスクリーニングにより同定され得る。第一のアプローチでは、宿主細胞中に挿入された先祖ウイルス配列核酸の存在は、挿入核酸と相同である配列を含むプローブを用いた核酸ハイブリダイゼーションにより検出され得る。

【0099】

第二のアプローチでは、発現ベクター／宿主系は、所望の核酸を含有するベクターの挿入により引き起こされるある種の「マーカー」遺伝子機能（例えばチミジンキナーゼ活性、抗生物質に対する耐性、形質転換表現型、バキュロウイルス中での閉塞体形成等）の存在または非存在に基づいて同定され、選択され得る。例えば、核酸がベクターのマーカー遺伝子配列内に挿入される場合、先祖ウイルス配列を含有する組換え体は、マーカー遺伝子機能の非存在により同定され得る。

【0100】

第三のアプローチでは、発現ベクター／宿主系は、組換え宿主生物により発現

される先祖ウイルス配列ポリペプチドに関して検定することにより同定され得る。このような検定は、例えば、*in vitro*検定系における先祖ウイルス配列ポリペプチドの物理的または機能的特性（例えば抗体による結合）を基礎にし得る。第四のアプローチでは、発現ベクター／宿主細胞は、既知の組換えDNA法により形質転換化宿主細胞をスクリーニングすることにより同定され得る。

【0101】

適切な発現ベクター宿主系および増殖状態が一旦確立されれば、当業界で既知の方法を用いてそれを増殖させ得る。さらに、挿入核酸配列の発現を調整する、あるいは所望の特定の方式で遺伝子産物を修飾またはプロセッシングする宿主細胞が選択され得る。ある種のプロモーターからの発現は、ある種の誘導物質の存在下で増大され得る；したがって先祖ウイルス配列の発現が制御され得る。さらに、ポリペプチドの翻訳および翻訳後プロセッシングおよび修飾（例えばグリコシル化またはリン酸化）のための特徴的および特異的メカニズムを有する異なる宿主細胞が用いられ得る。適切な細胞株または宿主系が選定されて、発現ポリペプチドの所望の修飾およびプロセッシングを保証する。例えば、細菌系における発現を用いて、非グリコシル化ポリペプチドを生成し得る。

【0102】

先祖タンパク質

本発明はさらに、確定先祖ウイルス配列を基礎にした先祖タンパク質に関する。このような先祖タンパク質としては、例えば全長タンパク質、ポリペプチド、その断片、誘導体および類似体が挙げられる。一局面において、本発明は先祖タンパク質のアミノ酸配列を提供する（例えば、表2および4；配列番号2；配列番号4参照）。いくつかの実施態様では、先祖タンパク質は機能的に活性である。先祖タンパク質、断片、誘導体および類似体は、典型的には所望の免疫原性または抗原性を有し、そして、例えばイムノアッセイに、免疫感作に、ワクチン等に用いられ得る。特定の実施態様は、抗体に結合され得る先祖タンパク質、断片、誘導体または類似体に関する。このような先祖タンパク質、断片、誘導体または類似体は、当業界で既知の手法により、所望の免疫原性に関して試験され得る（例えば、Harlow and Lane、上記、参照）。

【0103】

別の局面では、先祖タンパク質の少なくとも8~10連続アミノ酸を有する断片からなるかまたはそれを含むポリペプチドが提供される。他の実施態様では、断片は先祖タンパク質の少なくとも20または50連続アミノ酸を含む。その他の実施態様では、断片は35、100または200アミノ酸以下である。

【0104】

先祖タンパク質誘導体および類似体は、当業界で既知の種々の方法により製造され得る。それらの産生を生じる操作は、遺伝子またはタンパク質レベルで起こり得る。例えば先祖タンパク質をコードする核酸は、当業界で既知の多数の戦略のいずれかにより（例えば、Sambrook et al., 上記、参照）、例えば保存的置換、欠失、挿入等を作製することにより修飾され得る。核酸配列は、制限エンドヌクレアーゼ（単数または複数）を用いて適切な部位で切断され、その後さらに酵素修飾され、所望により単離され、そしてin vitroで結繋され得る。

【0105】

先祖タンパク質の断片、誘導体または類似体をコードする核酸の産生に際して、リーディングフレームが翻訳停止シグナルあるいは断片、誘導体または類似体の合成を妨げるその他のシグナルにより中断されないように、修飾化核酸は典型的には適正翻訳リーディングフレーム内に残存する。先祖ウイルス配列核酸は、in vitroまたはin vivoでも突然変異化されて、翻訳開始および／または終止配列を作製および／または破壊し得る。先祖ウイルス配列コード核酸は突然変異化されて、コード領域内に変異を作製し、および／または新規の制限エンドヌクレアーゼ部位を形成するかまたは既存のものを破壊し、そしてさらにin vitro修飾を促し得る。当業界で既知の突然変異誘発のための任意の技法が用いられ得るが、その例としては、化学的突然変異誘発、in vitro部位特異的突然変異誘発等が挙げられる（これらに限定されない）。

【0106】

先祖ウイルス配列の操作は、タンパク質レベルでなされ得る。本発明の範囲内に含まれるのは、合成中または合成後に区別的に修飾される（例えばin vivoまたはin vitro翻訳）先祖タンパク質断片、誘導体または類似体である。このよう

な修飾としては、保存的置換、グリコシル化、アセチル化、リン酸化、アミド化、既知の保護／遮断基による誘導化、タンパク質分解的切断、抗体分子またはその他の細胞性リガンドとの連結等が挙げられる。多数の化学的修飾のいずれかが、既知の技法により実行され、その例としては特異的化学切断（例えば臭化シアンによる）；酵素的切断（例えばトリプシン、キモトリプシン、パパイン、V8プロテアーゼ等による）；例えば NaBH_4 アセチル化、ホルミル化、酸化および還元による修飾；ツニカマイシンの存在下での代謝的合成等が挙げられる。

【0107】

さらに、先祖タンパク質の断片、誘導体および類似体は、化学的に合成され得る。例えば、所望のドメインを含む先祖タンパク質の一部または断片に対応するペプチドは、例えば自動ペプチド合成機を用いた化学的合成法の使用により合成され得る (Hunkapiller et al., Nature 310:105-11(1984); Stewart and Young, Solid Phase Peptide Synthesis, 2nd ed., Pierce Chemical Co., Rockford, IL, (1984)も参照)。さらに、所望により、非古典的アミノ酸または化学的アミノ酸類似体が、置換または付加としてポリペプチド配列中に導入され得る。

【0108】

非古典的アミノ酸としては、一般的アミノ酸のD-異性体、 α -アミノイソ酪酸、4-アミノ酪酸、2-アミノ酪酸、6-アミノヘキサン酸、2-アミノイソ酪酸、3-アミノプロピオン酸、オルニチン、ノルロイシン、ノルバリン、ヒドロキシプロリン、サルコシン、シトルリン、システイン酸、 ϵ -アチルグリシン、 ϵ -アチルアラニン、フェニルグリシン、シクロヘキシルアラニン、 β -アラニン、セロノシステイン、フルオロアミノ酸、デザイナアミノ酸、例えば β -メチルアミノ酸、 C_{α} -メチルアミノ酸、 N_{α} -メチルアミノ酸、ならびにその他のアミノ酸類似体が挙げられるが、これらに限定されない。さらにアミノ酸は、D（右旋性）またはL（左旋性）であり得る。

【0109】

先祖タンパク質、断片、誘導体または類似体は、そのアミノまたはカルボキシ末端で、ペプチド結合により異なるタンパク質のアミノ酸配列と連結される先祖タンパク質、その断片、誘導体または類似体（典型的には先祖タンパク質の少な

くともドメインまたはモチーフ、あるいは先祖タンパク質の少なくとも10連続アミノ酸からなる)を含むキメラまたは融合タンパク質でもあり得る。一実施態様では、このようなキメラタンパク質は、キメラタンパク質をコードする核酸の組換え体発現により産生される。キメラ核酸は、適切な核酸配列を適正リーディングフレーム内で互いに結繋し、そして当業界で一般に既知の方法によりキメラ産物を発現することにより作製され得る。あるいは、キメラタンパク質は、タンパク質合成技法により(例えば自動ペプチド合成機の使用により)製造され得る。

【0110】

先祖タンパク質は、標準的方法により、例えばクロマトグラフィー(例えば、イオン交換、アフィニティー、サイジングカラムクロマトグラフィー、高圧液体クロマトグラフィー)、遠心分離、示差溶解度により、またはタンパク質の精製のための任意のその他の標準技法により単離され、精製され得る。

【0111】

先祖タンパク質、断片、誘導体および類似体に対する抗体

先祖タンパク質(その断片、誘導体および類似体を含む)は、このような先祖タンパク質を、ならびに循環変異体と免疫特異的に結合する抗体を生成するための免疫原として用いられ得る。このような抗体としては、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖抗体、抗原結合抗体断片(例えばFab、Fab'、F(ab')₂、Fvまたは超可変部)ならびにFab発現ライブラリーが挙げられるが、これらに限定されない。いくつかの実施態様では、先祖タンパク質に対するポリクローナルおよび/またはモノクローナル抗体が産生される。他の実施態様では、先祖タンパク質のドメインに対する抗体が産生される。さらに他の実施態様では、免疫原性(例えば親水性)と同定される先祖タンパク質の断片は、抗体製造のための免疫原として用いられる。

【0112】

当業界で既知の種々の手法は、ポリクローナル抗体の製造のために用いられ得る。このような抗体の製造のためには、種々の宿主動物(例えばウサギ、マウス、ラット、ヒツジ、ヤギ、ラクダ等が挙げられるが、これらに限定されない)は

、先祖タンパク質、断片、誘導体または類似体を用いた注射により免疫感作され得る。宿主種によって種々のアジュバントを用いて免疫学的応答を増大し得るが、その例としては、フロイントアジュバント（完全および不完全）、無機ゲル、例えば水酸化亜アルミニウム、界面活性剤、例えばリソレシチン、プルロニックポリオール、多価陰イオン、ペプチド、オイルエマルション、カギアナカサガイヘモシアニン、ジニトロフェノール、ならびに有用であると考えられるヒトアジュバント、例えばBCG（カルメットゲラン桿菌）およびコリネバクテリウム属の *Corynebacterium parvum* が挙げられる（これらに限定されない）。

【0113】

先祖タンパク質、その断片、誘導体または類似体に向けられるモノクローナル抗体の調製に関しては、培養中の連続細胞株による抗体分子の産生を提供する任意の技法が用いられ得る。このような技法としては、例えばKohlerとMilstein（例えば、Nature 256:495-97(1975)参照）により最初に開発されたハイブリドーマ技法、トリオーマ技法（例えば、Hagiwara and Yuasa, Hum, Antibodies Hybridomas, 4:15-19(1993); Hering et al., Biomed. Biochim. Acta 47:211-16(1988)参照）、ヒトB細胞ハイブリドーマ技法（例えば、Kozbor et al., Immunology Today 4:72(1983)参照）、ならびにヒトモノクローナル抗体を産生するためのEBV-ハイブリドーマ技法（例えば、Cole et al., In: Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96(1985)参照）、が挙げられる。ヒト抗体が用いられ、そしてヒトハイブリドーマ（例えば、Cote et al., Proc. Natl. Acad. Sci. USA 80:2026-30(1983)参照）を用いることにより、またはin vitroでEBVウイルスでヒトB細胞を形質転換する（例えば、Cote et al., 上記、参照）ことにより得られる。

【0114】

さらに本発明のために、「キメラ」または「ヒト化」抗体（例えば、Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851-55(1984); Neuberger et al., Nature 312:604-08(1984); Takeda et al., Nature 314:452-54(1985)参照）が調製され得る。このようなキメラ抗体は、典型的には、適切な生物学的活性のヒト抗体分子からの遺伝子と一緒に、先祖タンパク質に特異的な抗体分子に関する非

ヒト遺伝子をスプライシングすることにより調製される。非ヒト抗体の抗原結合領域（例えばFab'、F(ab')₂、Fab、Fvまたは超可変部）を、実質的ヒト分子を産生するために組換えDNA技法によりヒト抗体の枠組み構造中に移すことが望ましい。

【0115】

このような「キメラ」分子の製造方法は一般的に周知であり、例えば、米国特許第4,816,567号、第4,816,397号、第5,693,762号および第5,712,120号；国際特許公告WO 87/02671およびWO 90/00616；ならびに欧州特許公告EP 239 400（これらの記載内容は、参照により本明細書中に含まれる）に記載されている。

【0116】

あるいはヒトモノクローナル抗体またはその一部は、Huse等（Science 246:1275-81(1989)）により一般的に記述された方法にしたがって先祖タンパク質と特異的に結合する抗体をコードするDNA分子に関して、先ずヒトB細胞cDNAライブラリーをスクリーニングすることにより同定され得る。次にDNA分子がクローン化され、所望の特異性を有する抗体（または結合ドメイン）をコードする配列を得るために増幅され得る。ファージ表示技法は、先祖タンパク質、その断片、誘導体または類似体と結合する抗体を選定するためのもうひとつの技法を提供する（例えば、国際特許公告WO 91/17271およびWO 92/01047；Huse et al., 上記、参照）。

【0117】

本発明の別の観点によれば、一本鎖抗体の産生に関して記載された技法（例えば、米国特許第4,946,778号および第5,969,108号参照）は、一本鎖抗体を産生するために適応され得る。本発明の付加的局面は、Fab発現ライブラリーの構築に関して記載された技法（例えば、Huse et al., 上記、参照）を利用して、先祖タンパク質、その断片、誘導体または類似体に対する所望の特異性を有するモノクローナルFab断片の迅速且つ容易な同定を可能にする。

【0118】

分子のイディオタイプを含有する抗体は、既知の技法により生成され得る。例えば、このような断片としては、抗体分子のペプシン消化により産生され得るFC

ab[']₂断片、F(ab['])₂断片のジスルフィド架橋を還元することにより生成され得るFab[']断片、パパインおよび還元剤で抗体分子を処理することにより生成され得るFab断片、ならびにFv断片が挙げられるが、これらに限定されない。組換えFv断片も、例えば米国特許第5,965,405号に記載された方法を用いて真核生物細胞中で産生され得る。

【0119】

抗体の産生において、所望の抗体に関するスクリーニングは、当業界で既知の技法（例えばELISA（酵素結合イムノソルベント検定））により成し遂げられ得る。一例において、先祖タンパク質の特異的ドメインを認識する抗体を用いて、そのドメインを含有するポリペプチドと結合する物質に関して生成されたハイブリドーマを検定し得る。先祖タンパク質のドメインに特異的な抗体も提供される。

【0120】

先祖タンパク質（断片、誘導体および類似体を含む）に対する抗体は、当業界で既知の方法にしたがって、受動抗体処理のために用いられ得る。抗体は、ウイルス感染を予防または治療するために個体に導入され得る。典型的にはこのような抗体療法は、ワクチン接種プロトコールに対するアジュバントとして実施される。抗体は前記と同様に製造され、そしてポリクローナルまたはモノクローナル抗体であり得るし、静脈内に、腸内に（例えば腸溶性被覆錠剤形態として）、エアロゾルにより、経口的に、経皮的に、経粘膜的に、胸膜腔内に、鞘内に、またはその他の適切な経路により投与され得る。

【0121】

免疫学的組成物およびワクチン

本発明は、免疫学的組成物、例えばワクチンも提供する。本発明の配列を用いるワクチン（「デジタルワクチン」）の開発の一例は、図4に示されている。本発明は、HIV先祖ウイルス配列（例えばHIVenvまたはgag遺伝子またはポリペプチド）を用いてワクチンを製造するための新しい方法も提供する。このような先祖ウイルス配列は、典型的には、実際の生物学的存在—創始ウイルス（即ち「ウイルスのイブ」）の構造に対応する。

【0122】

処方物

免疫原的有効量の1つまたはそれ以上の先祖ウイルスタンパク質配列あるいはその断片、誘導体または類似体を含む免疫原性組成物およびワクチンが提供される。先祖タンパク質配列中の免疫原性エピトープは当業界で既知の方法により同定され、それらのエピトープを含むタンパク質、断片、誘導体または類似体は、種々の方法により、ワクチン組成物中に送達され得る。適切な組成物としては、例えばリポペプチド（例えば、Vitiello et al., J. Clin. Invest. 95:341(1995)）、ポリ（DL-ラクチド-コグリコリド）（「PLG」）微小球中に封入されたペプチド組成物（例えば、Eldridge et al., Molec. Immunol. 28:287-94(1991); Alonso et al., Vaccine 12:299-306(1994); Jones et al., Vaccine 13:675-81(1995)参照）、

【0123】

免疫刺激複合体（ISCOMS）中に含入されたペプチド組成物（例えば、Takahashi et al., Nature 344:873-75(1990); Hu et al., Clin. Exp. Immunol. 113:235-43(1998)参照）、多重抗原ペプチド系（MAP）（例えば、Tam, Proc. Natl. Acad. Sci. USA. 85:5409-13(1988); Tam, J. Immunol. Methods 196:17-32(1996)参照）、ウイルス送達ベクター（例えば、Perkus et al., In: Concepts in vaccine development, Kaufmann(ed.), p.379(1996)参照）、ウイルスまたは合成期限の粒子（例えば、Kofler et al., J. Immunol. Methods. 192:25-35(1996);

【0124】

Eldridge et al., Sem. Hematol. 30: 16 (1993); Falo et al., Nature Med. 7:649(1995)参照）、アジュバント（例えば、Warren et al., Annu. Rev. Immunol. 4:369(1986); Gupta et al., Vaccine 11:293(1993)参照）、リポソーム（例えば、Reddy et al., J. Immunol. 148:1585(1992); Rock, Immunol. Today 17:131(1996)参照）、あるいは裸または粒子吸収cDNA（例えば、Shiver et al., In: Concepts in vaccine development, Kaufmann(ed.), p.423(1996)参照）が挙げられる。受容体媒介性ターゲティングとしても既知の毒素標的化送達技法、例えばAvant Immunotherapeutics, Inc. (Needham, Massachusetts)のものも

用いられ得る。

【0125】

さらに、本発明の免疫原性組成物およびワクチンとともに用いられ得る有用な担体は当業界で周知であり、その例としては、例えばチログロブリン、アルブミン、例えばヒト血清アルブミン、破傷風毒素、ポリアミノ酸、例えばポリL-リシン、ポリL-グルタミン酸、インフルエンザ、B型肝炎ウイルスコアタンパク質等が挙げられる。組成物およびワクチンは、生理学的耐容性（即ち許容可能な）希釈剤、例えば水または食塩水、典型的にはリン酸塩緩衝化生理食塩水を含有し得る。

【0126】

組成物およびワクチンは、典型的にはアジュバントも含む。アジュバント、例えば不完全フロイントアジュバント、リン酸アルミニウム、水酸化アルミニウムまたはミョウバンは、当業界で周知の物質の例である。さらに、本明細書中に開示されているように、CTL応答は、先祖タンパク質（またはその断片、誘導体または類似体）を脂質、例えばトリパルミトイル-S-グリセリルシステイニル-セリル-セリン（P₃CSS）と接合させることにより用意され得る。

【0127】

本明細書中にさらに詳細に開示されているように、注射、エロゾル、経口、経皮、経粘膜、胸膜腔内、鞘内またはその他の適切な経路により、本発明にしたがって先祖ウイルス配列タンパク質組成物を含有する組成物またはワクチンで免疫感作する場合、宿主の免疫系は、大量のCTL、HTLおよび／または所望の抗原に特異的な抗体を産生することにより組成物またはワクチンに応答する。したがって宿主は、典型的には、後の感染に対して少なくとも部分的に免疫になるか、または進行中の慢性感染の発症に対して少なくとも部分的に耐性になるか、あるいは少なくとも何らかの治療利益を得る。

【0128】

治療的または予防的免疫感作に関しては、先祖タンパク質（断片、誘導体および類似体を含む）は、ウイルスまたは細菌ベクターによっても発現され得る。発現ベクターの例としては、弱毒化ウイルス宿主、例えばワクシニアまたは鶏痘が

挙げられる。一実施態様では、このアプローチは、例えばポリペプチドをコードするヌクレオチド配列を発現するためのベクターとしてのワクシニアウイルスの使用を包含する。急性または慢性感染宿主中への、あるいは非感染宿主中への導入時に、組換えワクシニアウイルスは免疫原性タンパク質を発現し、それにより宿主CTL、HTLおよび/または抗体応答を引き出す。

【0129】

免疫感作プロトコールにおいて有用なワクシニアベクターおよび方法は、例えば米国特許第4,722,848号（この開示内容は、参照により本明細書中に含まれる）に記載されている。本発明のペプチドの治療的投与または免疫感作に有用な広範な種々のその他のベクター、例えばアデノおよびアデノ随伴ウイルスベクター、レトロウイルスベクター、ネズミチフス菌ベクター、無毒化炭疽菌毒素ベクター、アルファウイルス等も、本明細書中の説明から当業者に明らかになるように、用いられ得る。用いられ得るアルファウイルスベクターとしては、例えばシンジビスおよびベネズエラウマ脳脊髄炎（VEE）ウイルスが挙げられる（例えば、Coppola et al., J. Gen. Virol. 76:635-41(1995); Caley et al., Vaccine 17:3124-35(1999); Loktev et al., J. Biotechnol. 44:129-37(1996)参照）。

【0130】

1つまたはそれ以上の先祖タンパク質（断片、誘導体または類似体を含む）をコードするポリヌクレオチド（例えばDNAまたはRNA）も、患者に投与され得る。このアプローチは、例えばWolff等（Science 247:1465(1990)）に、米国特許第5,580,859号、第5,589,466号、第5,804,566号、第5,739,118号、第5,736,524号、第5,679,647号およびWO 98/04720に、そして以下で詳細に記載されている。DNAベースの送達技法の例としては、「裸DNA」、促進（プピジカイン、ポリマーまたはペプチド媒介性）送達、陽イオン性脂質複合体、粒子媒介性（「遺伝子銃」）または圧力媒介性送達（例えば、米国特許第5,922,687号参照）が挙げられる。

【0131】

ワクチン接種の一手段としてのタンパク質抗原をコードする裸プラスミドDNAの直接注入は、過去10年間に開発されたいくつかのHIV送達および発現系の中で

、大いに注意を引くものである。マウスモデルでは、ならびに大型動物モデルにおいては、体液性および細胞性の両免疫応答が容易に誘導されて、いくつかの場合には試験感染に対する防御免疫を生じる。セムリキ森林熱ウイルス (SFV) レプリコンも、例えば裸DNA免疫感作の状況において用いられ得る。SFVは、アルファウイルス科に属し、この場合ゲノムは、それ自体のレプリカーゼをコードする正の極性を有する一本鎖RNAからなる。

【0132】

SFV構造遺伝子を当該遺伝子に取り替えることにより、総細胞タンパク質の25%という高い発現レベルが得られる。プラスミドベクターを上回るこのアルファウイルスのもうひとつの利点は、その非持続性である。当該抗原は、高レベルで、しかし短期間（典型的には<72時間）発現される。これに対比して、プラスミドベクターは一般に、長時間に亘って当該抗原の合成を誘導し、外来DNAおよび細胞形質転換の染色体組込みを危険に曝す。さらに抗原持続性または少量の抗原の反復接種は、寛容を誘導することが実験的に示されている。したがって長期抗原合成は、理論的に、免疫よりむしろ非応答性を生じ得る。

【0133】

先祖タンパク質、断片、誘導体および類似体は、*in vivo*または*ex vivo*でも被験者に導入され得る。例えば先祖ウイルス配列は、限定細胞集団中に導入され得る。遺伝子導入のための適切な方法としては、例えば以下のものが挙げられる：

- 1) 直接遺伝子導入（例えば、Wolff et al., Science 247: 1465-68(1990)参照）。
- 2) リポソーム媒介性DNA導入（例えば、Caplen et al., Nature Med. 3:39-46(1995); Crystal, Nature Med. 1:15-17(1995); Gao and Huang, Biochem. Biophys. Res. Comm. 179:280-85(1991)参照）。

【0134】

- 3) レトロウイルス媒介性DNA導入（例えば、Kay et al., Science 262:117-19(1993); Anderson, Science 256:808-13(1992)参照）。レトロウイルスプラスミドベクターが得られるレトロウイルスとしては、レンチウイルスが挙げられる。それらの例としてはさらに、モロニーネズミ白血病ウイルス、脾臓壊死ウイルス

、レトロウイルス、例えばラウス肉腫ウイルス、ハーベイ肉腫ウイルス、鳥類白血症ウイルス、テナガザル白血病ウイルス、ヒト免疫不全ウイルス、骨髄増殖性肉腫ウイルスおよび乳癌ウイルスが挙げられるが、これらに限定されない。

【0135】

一実施態様では、レトロウイルスプラスミドベクターは、モロニーネズミ白血病ウイルスから得られる。遺伝子両方におけるレトロウイルスベクターの使用を示す例としては、以下のものが挙げられる：Clowes et al.(J. Clin. Invest. 93:644-51(1994)；Kiem et al.(Blood 83:1467-73(1994)；Salmons and Gunzberg (Human Gene Therapy 4:129-41(1993)；およびGrossman and Wilson(Curr. Opin. In Genetics and Devel. 3:110-14(1993))。

【0136】

4) DNAウイルス媒介性DNA導入。このようなDNAウイルスとしては、アデノウイルス（例えばAd-2またはAd-5ベースのベクター）、ヘルペスウイルス（典型的には単純ヘルペスウイルスベースのベクター）およびパルボウイルス（例えば「欠陥」または非自律性パルボウイルスベースのベクターまたはアデノ随伴ウイルスベースのベクター、例えばAAV-2ベースのベクター）（例えば、Ali et al., Gene Therapy 1:367-84(1994)；米国特許第4,797,368号および第5,139,941号参照）（これらの開示内容は、参照により本明細書中に含まれる）。

【0137】

アデノウイルスは、広範な宿主範囲を有し、静止性または末期的分化細胞、例えばニューロンまたは肝細胞を感染し得るし、本質的に非癌性と思われる。アデノウイルスは、宿主ゲノム中に組み込まれるとは思えない。それらは染色体外に存在するため、挿入性突然変異誘発の危険は大いに低減される。アデノ随伴ウイルスは、アデノウイルスベースのベクターと同様の利点を示す。しかしながらAAVは、ヒト第19染色体上での部位特異的組込みを示す。

【0138】

KozarskyとWilson (Current Opinion in Genetics and Development 3:499-503(1993)) は、アデノウイルスベースの遺伝子療法の再検討を示す。Bout等 (Human Gene Therapy 5:3-10(1994)) は、アカゲザルの呼吸上皮に遺伝子を導入する

ためのアデノウイルスベクターの使用を実証する。Herman等 (Human Gene Therapy 10:1239-49(1999)) は、段階I臨床試験におけるヒト前立腺中への単純ヘルペスチミジンキナーゼ遺伝子を含有する複製欠陥アデノウイルスの前立腺内注入と、その後のプロドラッグガンシクロビルの静脈内投与を記載する。遺伝子療法におけるアデノウイルスの使用のその他の例は、Rosenfeld等 (Science 252:431-34(1991)) ; Rosenfeld等 (Cell 68:143-55(1992)) ; Mastrangeli等 (J. Clin. Invest. 91:225-34(1993)) ; Thompson (Oncol. Res. 11:1-8(1999)) に見出され得る。

【0139】

当該先祖ウイルス配列を導入するための特定のベクター系の選定は、種々の因子によっている。重要な一因子は、標的細胞集団の性質である。レトロウイルスベクターは広範に研究され、そして多数の遺伝子療法用途に用いられてきたが、しかしこれらのベクターは一般に、非分裂中細胞を感染するには適していない。さらにレトロウイルスは、腫瘍原性に対する潜在性を有する。しかしながらレンチウイルスベクターの分野における近年の発達は、これらの制限のいくつかを阻止し得る (Naldini et al., Science 272:263-67(1996)参照)。

【0140】

先祖タンパク質あるいはその断片、誘導体または類似体をコードする核酸を含有する任意の適切な発現ベクターは本発明にしたがって用いられ得る、と当業者は理解する。このようなベクターを構築するための技術は既知である (例えば、Anderson, Nature 392:25-30(1998); Verma, Nature 389:239-42(1998)参照)。標的部位へのベクターの導入は、既知の技法を用いて成し遂げられ得る。

【0141】

もうひとつの実施態様では、DNA導入ベクター (pJW4304SV40/EBVベクター (pJW4304SV40/EBVベクターは、Robinson et al., Ann. New York Acad. Sci. 27:209-11(1995)およびYasutomi et al., J. Virol. 70: 678-81 (1996)により記載されているpJW4303から調製された)) を用いる新規発現系は、極高効率性RNA/タンパク質発現系 (セムリキ森林熱ウイルス) とともに、安全且つ安価なワクチンを用いてワクチン接種された宿主細胞中での最大タンパク質発現を達成するた

めに用いられる。SFVcDNAは、例えばサイトメガロウイルス (CMV) プロモーターの制御下に置かれる (図7参照)。慣用的DNAベクターと違って、CMVプロモーターは、抗原コード核酸の発現を直接駆動しない。

【0142】

代わりにそれは、組換えSFVレプリコンRNA転写体の合成を指図する。このRNA分子の翻訳は、組換えRNAの細胞質自己増幅を触媒するSFVレプリカーゼ複合体を生成し、その結果として実際の抗原コードmRNAを高レベルで生成する。ベクター送達後、トランスフェクト化宿主細胞は2〜3日以内に死ぬ。本発明の状況では、envおよび/またはgag遺伝子は典型的にはこのベクター中でクローン化される。ノーザンブロット、ウエスタンブロット、SDS-PAGE、免疫沈降検定およびCD4結合検定を用いる in vitro 実験は、タンパク質発現レベル、タンパク質特性、発現持続時間およびベクターの細胞障害作用を査定することによりこの系の効率を確定するために、下記のように実施され得る。

【0143】

いくつかの実施態様では、先祖タンパク質 (その断片、誘導体または類似体を含む) は、それを必要とする被験者に投与される。初期治療的免疫感作のための投与量は一般に、低いほうの値が約1、5、50、500または1,000 μ gおよび高いほうの値が10,000、20,000、30,000または50,000 μ gである単位投与量範囲で生じる。ヒトに関する投与量値は、典型的には約500 μ g〜50,000 μ g/70 kg患者の範囲である。数週間〜数ヶ月に亘る増強レジメンに準じた約1.0 μ g〜約50,000 μ gのポリペプチドの増強投与量は、患者の血液から得られたCTLおよびHTLの抗体レベルまたは特異的活性を測定することにより確定した場合の患者の応答および症状によって投与され得る。

【0144】

ヒト単位用量形態のタンパク質または核酸組成物は、典型的には、ヒト単位用量の許容可能担体、典型的には水性担体を含む製剤組成物中に含入され、そしてヒトへのこのような組成物の投与のために用いられることが当業者に既知である流体容積中に投与される (例えば、Remington "Pharmaceutical Sciences", 17 Ed., Gennaro (ed.), Mack Publishing Co., Easton, Pennsylvania(1985)参

照)。

【0145】

先祖タンパク質および核酸は、リボソームを介しても投与され、これは特定組織、例えばリンパ組織に対してペプチドを標的化するのに、あるいは感染細胞に対して選択的に標的化するのに、ならびに組成物の半減期を増大するのに役立つ。リボソームとしては、エマルション、発泡体、ミセル、不溶性単一層、液体結晶、リン脂質分散液、薄板層等が挙げられる。これらの調製において、送達されるタンパク質または核酸は、単独でまたはリンパ細胞の間で流布している受容体と結合する分子、例えばCD45抗原と結合するモノクローナル抗体と一緒に、あるいはその他の治療用または免疫原性組成物と一緒に、リボソームの一部として組み入れられる。

【0146】

したがって、所望のタンパク質または核酸で満たされたまたは装飾されたりボソームは、リンパ細胞の部位に向けられ、そこでは、リボソームは次に、タンパク質組成物を細胞に送達する。本発明にしたがって用いるためのリボソームは、一般的に中性および陰性荷電リン脂質およびステロール、例えばコレステロールを含む標準小胞形成脂質から生成される。脂質の選定は、一般的には、例えばリボソームサイズ、酸不安定性および血流中のリボソームの安定性の考慮により導き出される。種々の方法は、例えばSzoka et al., Ann. Rev. Biophys. Bioeng. 9:467(1980)ならびに米国特許第4,235,871号、第4,501,728号、第4,837,028号および第5,091,369号に記載されたように、リボソームを調製するために利用可能である。

【0147】

免疫系の細胞をターゲッティングするために、リボソーム中に組み入れられるリガンドとしては、例えば所望の免疫系細胞の細胞表面決定基に特異的な抗体またはその断片が挙げられ得る。タンパク質または核酸を含有するリボソーム懸濁液は、例えば静脈内に、局部的に、局所的等で、とりわけ投与方法、送達中のタンパク質または核酸等によって変わる用量で、投与され得る。

【0148】

固体組成物に関しては、慣用的非毒性固体担体が用いられ、その例としては例えば医薬品等級のマニトール、ラクトース、デンプン、ステアリン酸マグネシウム、ナトリウムサッカリン、タルク、セルロース、グルコース、スクロース、炭酸マグネシウム等が挙げられる。経口投与に関しては、通常用いられる賦形剤、例えば前記の担体、および一般的に10~95%の、典型的には25%~75%の濃度での活性成分、即ち先祖タンパク質または核酸のいずれかを組入れることにより、製薬上許容可能な非毒性組成物が生成される。

【0149】

エーロゾル投与に関しては、免疫原性タンパク質または核酸は、典型的には微粉碎形態で、界面活性剤および噴射剤とともに存在する。ペプチドの適切なパーセンテージは約0.01重量%~20重量%、典型的には約1重量%~10重量%である。界面活性剤は、もちろん非毒性であり、そして典型的には噴射剤中で可溶性である。このような作用物質の代表は、約6~22個の炭素原子を含有する脂肪酸のエステルまたは部分エステル、例えばカプロン酸、オクタン酸、ラウリン酸、パルミチン酸、ステアリン酸、リノール酸、リノレン酸、ステアリン酸およびオレイン酸であって、脂肪族多価アルコールまたはその環状無水物を伴う。混合エステル、例えば混合または天然グリセリドが用いられ得る。界面活性剤は、組成物の約0.1重量%~約20重量%、典型的には0.25~5%を構成し得る。組成物の残余は、普通は噴射剤である。例えば鼻腔内送達のためにレシチンを用いるように、所望により担体も含入され得る。

【0150】

先祖ウイルス配列により引き出される免疫応答

先祖タンパク質（断片、誘導体および類似体を含む）は、上記のようにワクチンとして用いられ得る。「デジタルワクチン」と呼ばれるこのようなワクチンは、典型的には、既存ウイルスの任意の配列によりまたはコンセンサス配列によりコードされるタンパク質抗原を含むワクチンより大きい循環系統の分画に対する中和抗体および／またはウイルス（例えばHIV）特異的CTLを引き出すものに関してスクリーニングされる。このようなデジタルワクチンは、典型的には、先祖ウイルス配列が得られたサブタイプと同一のウイルス（例えばHIV-1ウイルス）の

サブタイプにより試験された場合に、防御を提供する。

【0151】

本発明は、先祖ウイルス遺伝子配列の機能を分析するための方法も提供する。例えば一実施態様では、gp160先祖ウイルス遺伝子配列は、例えばCD4結合、補助受容体結合、受容体特異性（例えばCCR5受容体との結合）、タンパク質構造および細胞融合を引き起こす能力といった機能に関する検定により分析される。先祖配列は成育可能ウイルスを生じ得るが、しかしこのような成育可能ウイルスは好結果のワクチンを売るためには必要でない。例えば正確に折りたたまれないgp160先祖は、普通は免疫系に埋もれさせるエピトープを曝露することにより、より免疫原性になり得る。さらに先祖ウイルス配列はワクチンとして首尾よく用いられ得るが、しかしこのような配列は、免疫原（例えばワクチン）として用いられる場合、tatまたはrevのようなタンパク質をコードする代替的オープンリーディングフレームを含む必要がない。

【0152】

したがって一つの観点では、マウスは先祖タンパク質で免疫感作され、体液性および細胞性免疫応答に関して試験される。典型的には、5～10匹のマウスが、例えば50 μ l容量中に先祖ウイルス配列をコードするgagおよび/またはenv遺伝子を含有するプラスミドを皮内または筋肉注射される。2つの対照群は、典型的にはその結果を説明するために用いられる。一対照群は、標準実験室系統（例えばHIV-1-IIIIB）からのgagまたはenv遺伝子を含有する同一ベクターを注射される。第二対照群は、いかなる挿入物も有さない同一ベクターを注射される。

【0153】

Gagまたはenvタンパク質に対する抗体滴定は、下記のように、標準イムノアッセイ（例えばELISA）を用いて実施される。中和抗体は、サブタイプ特異的実験室HIV-1系統、例えばpNL4-3（HIV-1-IIIIB）、ならびにHIV-1感染個体からの一次単離物により分析される。広範な一次単離物を中和する先祖ウイルス配列タンパク質誘導中和抗体の能力は、免疫原性またはワクチン組成物を示す一因子である。同様の研究は、大型動物、例えば非ヒト動物（例えばマッカク属サル）またはヒトで実施され得る。

【0154】

先祖タンパク質誘導抗体を滴定するためのイムノアッセイ

試料中の抗体を検出するための当業者に既知の種々の検定がある（例えば、Harlow and Lane, 上記、参照）。概して、先祖タンパク質ワクチンで免疫感作された被験者中の抗体の存在または非存在は、（a）免疫感作被験者から得られた生物学的試料を1つまたはそれ以上の先祖タンパク質（その断片、誘導体または類似体を含む）と接触させ；（b）先祖タンパク質（単数または複数）と結合する抗体のレベルを試料中で検出し；そして（c）抗体のレベルを予定カットオフ値と比較することにより確定され得る。

【0155】

典型的実施態様では、検定は、試料からの抗体と結合しそして除去するために固体支持体上に固定された先祖タンパク質（断片、誘導体または類似体を含む）の使用を包含する。結合抗体は次に、レポーター基を含有する検出試薬を用いて検出され得る。適切な検出試薬としては、抗体／先祖タンパク質複合体およびレポーター基で標識された遊離タンパク質と結合する抗体を含む（例えば半競合検定において）。あるいは、当該先祖タンパク質と結合する抗体がレポーター基で標識され、抗原を試料とともにインキュベートした後に免疫感作抗原との結合を可能にされる競合検定が利用され得る。試料の構成成分が当該先祖タンパク質との標識化抗体の結合を抑制する程度は、免疫感作か先祖タンパク質との試料の反応性を示す。

【0156】

抗原が取り付けられ得る固体支持体は、当業者に既知の任意の固体物質であり得る。例えば固体支持体は、微小滴定プレート中の試験ウェル、あるいはニトロセルロースまたはその他の適切な膜であり得る。あるいは支持体は、ビーズまたは円板、例えばガラス、ガラス繊維、ラテックスまたはプラスチック材料、例えばポリスチレンまたはポリビニルクロリドであり得る。支持体は、磁気粒子または繊維光学センサー、例えば米国特許第5,359,681号（この記載内容は、参照により本明細書中に含まれる）に開示されているものであり得る。

【0157】

先祖タンパク質は、当業者に既知の種々の技法を用いて固体支持体に結合され得る（これは特許および科学文献中に詳細に記載されている）。本発明の状況において、「結合された」という用語は、非共有結合、例えば吸着、ならびに共有結合の両方を指す（例えば、Pierce Immunotechnology Catalog and Handbook, at A12-A13(1991)参照）。

【0158】

ある種の実施態様では、検定は酵素結合イムノソルベント検定（ELISA）である。この検定は、当該先祖タンパク質を認識する試料内に存在する抗体が固定タンパク質と結合されるように、まず、固体支持体上に、一般的には微小滴定プレートの上セル上に固定された先祖タンパク質を試料と接触させることにより実施され得る。次に非結合試料を固定先祖タンパク質から除去し、固定抗体-タンパク質複合体と結合し得る検出試薬が付加される。次に固体支持体に結合されたままの検出試薬の量が、特定の検出試薬に関して適切な方法を用いて確定される。

【0159】

特に、先祖タンパク質が前記のような支持体上に固定されれば、支持体上の残りのタンパク質結合部位は、典型的には遮断される。当業者に既知の任意の適切な遮断剤、例えばウシ血清アルブミンまたはトゥイーン™ 20 (Sigma Chemical Co., St Louis, Mo) が用いられ得る。次に固定先祖タンパク質は試料とともにインキュベートされ、抗体がタンパク質と結合される。試料は、インキュベーション前に適切な希釈剤、例えばリン酸塩緩衝化生理食塩水（PBS）で希釈され得る。概して、適切な接触時間（即ちインキュベーション時間）は、免疫感作被験者の生物学的試料内の抗体の存在を検出するのに十分な時間である。平衡を達成するのに必要な時間は、時間中に起こる結合のレベルを検定することにより容易に確定され得る、と当業者は理解する。室温では、約30分のインキュベーション時間が一般的に十分である。

【0160】

非結合試料は、適切な緩衝液、例えば0.1%トゥイーン™ 20を含有するPBSで固体支持体を洗浄することにより除去され得る。次に検出試薬が固体支持体に付加され得る。適切な検出試薬は、固定抗体-タンパク質複合体と結合し、そして当

業者に既知の種々の手段のいずれかにより検出され得る任意の化合物である。典型的には、検出試薬は、レポーター基に接合された結合剤（例えばプロテインA、プロテインG、免疫グロブリン、レクチンまたは遊離抗原）を含有する。

【0161】

適切なレポーター基としては、酵素（例えばホースラディッシュペルオキシダーゼまたはアルカリ性ホスファターゼ）、基質、補因子、阻害剤、染料、放射性核種、発光基、蛍光基およびビオチンが挙げられる。結合剤とレポーター基との接合は、当業者に既知の標準方法を用いて達成され得る。種々のレポーター基に予備接合される一般結合剤は、多数の商業的供給元（例えばZymed Laboratories, San Francisco, CAおよびPierce, Rockford, IL）から購入され得る。

【0162】

次に検出試薬が、結合抗体を検出するのに十分な長さの時間、固定抗体-タンパク質複合体とともにインキュベートされる。時間の適切な長さは、一般的にメーカーの使用説明書から、あるいは時間中に起こる結合のレベルを検定することにより確定され得る。次に非結合検定試薬が除去され、レポーター基を用いて結合検定試薬が検出される。レポーター基を検出するために用いられる方法は、レポーター基の性質によって異なる。放射性基に関しては、シンチレーション計数法またはオートラジオグラフィ法が一般的に適用している。分光分析法は、染料、発光基および蛍光基を検出するために用いられ得る。ビオチンは、異なるレポーター基（一般に、放射性または蛍光基または酵素）に結合されたアビジンを用いて検出され得る。酵素レポーター基は一般に、基質を付加し（一般的に特定時間）、その後反応産物の分光分析またはその他の分析を実施することにより検出され得る。

【0163】

試料中の抗先祖タンパク質抗体の存在または非存在を確定するために、固体支持体に依然として結合されるレポーター基から検出されたシグナルは、一般的に、予定カットオフ値に対応するシグナルと比較される。一実施態様では、カットオフ値は、固定先祖タンパク質が非免疫感作被験者からの試料とともにインキュベートされた場合に得られる平均シグナルである。

【0164】

関連実施態様では、検定は迅速フロースルーまたはストリップ試験フォーマットで実施され、この場合、先祖タンパク質は膜、例えばニトロセルロース、ナイロン、PVDF等の上に固定される。フロースルー試験では、試料内の抗体は、試料が膜を通過すると、固定ポリペプチドと結合する。検出試薬（例えばプロテインA-コロイド筋）は次に、検出試薬を含有する溶液が膜を通して流動すると、抗体-タンパク質複合体に結合する。次に結合検出試薬の検出は、前記のように実施され得る。ストリップ試験フォーマットでは、先祖タンパク質が結合される膜の一端は、試料を含有する溶液中に浸漬される。試料は検出試薬を含有する領域を通して固定先祖タンパク質の領域に、膜に沿って移動する。

【0165】

タンパク質での検出試薬の濃度は、試料中の抗先祖タンパク質抗体の存在を示す。典型的には、その部位での検出試薬の濃度は、視覚的に読取られ得る線のようなパターンを生じる。このようなパターンの非存在は、陰性結果を示す。概して、膜上に固定されるタンパク質の量は、生物学的試料が上記のように陽性シグナル（例えばELISAにおいて）を生成するのに十分である抗体のレベルを含有する場合、視覚的に識別可能なパターンを生じるよう選定される。典型的には、膜上に固定されるタンパク質の量は、約25 ng～約1 μ g、さらに典型的には約50 ng～約500 ngの範囲である。このような試験は、典型的には極少量（例えば1滴）の被験者血清または血液を用いて実施され得る。

【0166】

細胞傷害性Tリンパ球検定

HIV-1感染の治療における別の因子は、細胞性免疫応答、特にCD8⁺細胞傷害性Tリンパ球（CTL）を包含する細胞性免疫応答である。細胞傷害性Tリンパ球検定は、標準方法を用いて前記と同様に、同種および異種HIV系統に対する先祖ウイルス配列を用いたサブゲノム免疫感作後に細胞性免疫応答をモニタリングするため用いられ得る（例えば、Burke et al., 上記； Tigges et al. 上記、参照）。

【0167】

T細胞応答を検出するために利用される慣用的検定としては、例えば増殖検定

、リンホカイン分泌検定、直接細胞傷害性検定、限定希釈検定等が挙げられる。例えば、先祖タンパク質とともにインキュベートされた抗原提示細胞は、応答者細胞集団においてCTL応答を誘導する能力に関して検定され得る。抗原提示細胞は、末梢血単球または樹状細胞のような細胞であり得る。あるいは、内部プロセッシングペプチドをクラスI分子に負荷するそれらの能力を欠く、そして適切なヒトクラスI遺伝子でトランスフェクトされた突然変異体非ヒト哺乳類細胞株を用いて、*in vitro*一次CTL応答を誘導する当該先祖ペプチドの能力を試験し得る。

【0168】

末梢血単球 (PBMC) は、CTL前駆体の応答者細胞供給源として用いられ得る。適切な抗原提示細胞は、先祖タンパク質とともにインキュベートされ、その後、タンパク質負荷抗原提示細胞が、最適培養条件下で、応答者細胞集団とともにインキュベートされる。陽性CTL活性化は、放射能標識化標的細胞、即ち特異的ペプチドパルス化標的ならびにペプチド配列が得られた内因的プロセッシング形態の抗原を発現する標的細胞の両方を殺害するCTLの存在に関して培養を検定することにより確定され得る。

【0169】

別の適切な方法は、フルオレセイン標識化HLA四量体複合体を用いて染色することにより、抗原特異的T細胞の直接定量を可能にする (Altman et al., *Proc. Natl. Acad. Sci. USA* 90:10330(1993); Altman et al., *Science* 274:94(1996))。その他の相対的に新しい技術開発としては、細胞内リンホカインのための染色、ならびにインターフェロン放出検定またはELISPOT検定が挙げられる。四量体染色、細胞内リンホカイン染色およびELISPOT検定は、典型的には、慣用的検定より少なくとも10倍感受性である (Lalvani et al., *J. Exp. Med.* 186:859(1997); Dunbar et al., *Curr. Biol.* 8:413(1998); Murali-Krishna et al., *Immunity* 8:177(1998))。

【0170】

診断

本発明は、本明細書中に記載した先祖ウイルス配列を用いたウイルス (例えば

HIV) 感染および／またはAIDSの診断方法も提供する。ウイルス（例えばHIV）感染および／またはAIDSの診断は、当業者に周知の種々の方法を用いて実行され得る。このような方法としては、上記のようなイムノアッセイ、ならびに核酸配列の存在を検出するための組換えDNA法が挙げられるが、これらに限定されない。

【0171】

ウイルス遺伝子配列の存在は、例えば標準技法を用いて、表1または3に記述された配列またはその一部を用いて設計された特定のプライマーを用いたポリメラーゼ連鎖反応（PCR）により検出され得る（例えば、Innis et al., PCR Protocols A Guide to Methods and Application(1990); 米国特許第4,683,202号、第4,683,195号および第4,889,818号; Gyllensten et al., Proc. Natl. Acad. Sci. U.S.A. 85: 7652-56 (1988); Ochman et al., Genetics 120:621-23(1988); Loh et al., Science 243:217-20(1989)参照）。あるいは、ウイルス遺伝子配列は、当業者に周知の方法により、表1または3に記述された配列と少なくとも70%の同一性を有する核酸プローブを用いるハイブリダイゼーション法を用いて生物学的試料中で検出され得る（例えば、Sambrook et al., 上記、参照）。

【0172】

実施例

実施例1. 先祖ウイルス配列の確定

HIV-1サブタイプCの遺伝子を表示する配列を、GenBankおよびLos Alamos配列データベースから選択した。39のサブタイプC配列を用いた。18の外群配列（他の群Mサブタイプの各々からの2つ（図8））は、サブタイプC配列を捜し出すための外群として用いた。CLUSTALW (Thompson et al., Nucleic Acids Res. 22:4673-80(1994)) を用いて配列を整列させ、GDE (Smith et al., CABIOS 10: 671-5 (1994)) を用いてアラインメントを精製し、そしてアミノ酸配列をそれらから翻訳した。ギャップを、それらがコドン間に挿入されるように操作した。このアラインメント（アラインメントI）を、明瞭に整列され得ない領域を除去するように、系統発生分析用に修飾し (Learn et al., J. Virol. 70:5720-30(1996))、アラインメントIIを生じた。

【0173】

これらの配列（アラインメントII）に関する系統発生および先祖状態再構築のための適切な進化モデルを、Modeltest 3.0 (Posada and Crandall, Bioinformatics 14:817-8(1998)) で実行される場合、Akaike情報判定基準 (AIC) (Akaike, IEEE Trans. Autom. Contr. 19: 716-23 (1974)) を用いて選定した。サブタイプC先祖配列のための分析に関しては、最適モデルは、転写の両クラスに関しては等速度であり、トランスバージョンの4つのクラス全てに関しては異なる速度であって、不変部位、ならびに可変部位の部位-部位率変動性のX分布を伴う (TM+I+Gモデルとして言及される)。この場合のモデルのパラメーターを以下に示す：平衡ヌクレオチド頻度： $f_A=0.3576$ 、 $f_C=0.1829$ 、 $f_G=0.2314$ 、 $f_T=0.2290$ ；不変部位の割合=0.2447；X分布の形状パラメーター (α) =0.7623；速度マトリックス (R) マトリックス値： $R_{A \rightarrow C}=1.7502$ 、 $R_{A \rightarrow G}=R_{C \rightarrow T}=4.1332$ 、 $R_{A \rightarrow T}=0.6825$ 、 $R_{C \rightarrow G}=0.6549$ 、 $R_{G \rightarrow T}=1$ 。

【0174】

配列（アラインメントII）に関する進化系統樹は、PAUP* バージョン4.0b (Swofford, PAUP 4.0: Phylogenetic Analysis Using Parsimony (And Other Methods); Sinauer Associates, Inc. (2000)) で実行される場合、最大見込み概算 (MLE) 法を用いて推定した。特にサブタイプC配列に関しては、10の異なる亜系統樹-剪定-再移植 (SPR) 発見的検索を、各々異なる無作為付加順を用いて実施した。10のすべての検索が、同一MLE系統発生を見出した ($\ln L=-33585.74$)。サブタイプCに関する先祖ヌクレオチド配列を、この系統発生、データベース（アラインメントII）、ならびに限界見込み概算（下記参照）を用いた前記のTM+I+Gモデルを用いて、このサブタイプの基礎節点での配列であると推定した。

【0175】

この推定配列は、いくつかの可変部 (V1、V2、V4およびV5)、ならびに明瞭に整列され得ない4つの付加的短領域の部分に関する予測先祖配列を含まない（これら8つの領域は、アラインメントIIを生成するためにアラインメントIから除去される）。以下の手法を用いて、高可変部を含めた完全gp160に関するアミノ酸配列を予測した。推定先祖配列は視覚的にアラインメントIと一列に整列されて、GDE (Smith et al., 上記) を用いて翻訳される。高可変部を完全コドンとし

て欠失させたため、翻訳は正しいリーディングフレーム内であり、コドンは適正に維持された。

【0176】

アラインメントIIから欠失された領域に関する先祖アミノ酸配列を視覚的に予測し、コンピュータプログラムMacCladeバージョン3.08_a (Maddison and Maddison, MacClade-Analysis of Phylogeny and Character Evolution -Version 3. Sinauer Associates, Inc.(1992)) を用いたこれらの部位に関する最節約ベースの配列再構築を用いて精製した。このアミノ酸配列を、Wisconsin配列分析パッケージ (GCG) 、バージョン10のBACKTRANSLATEプログラムならびにコドン使用データベース ([http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Homo+sapiens+\[gbpri\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Homo+sapiens+[gbpri])) からのヒト遺伝子コドン表を用いて、ヒト細胞中での発現のために最適化されたDNA配列に転換した。

【0177】

実施例2.

所定のサブタイプに関する最大見込み系統発生を確定するためには、異なる方法が利用可能である。このような一方法は、大型進化手段から得た遺伝子配列の試料の系図の数学的説明である合一理論に基づいている。合一分析は、*in vivo* のならびにより大流行中のHIV集団を考慮に入れ、そして異なる過程をHIV集団に作用させた場合に、サンプリングした系図がどのように振舞うかを理解する一方法を提供する。この理論を用いて、先祖ウイルス配列、例えば始祖またはMRCAの配列を確定し得る。指数関数的増殖集団は、時間を遡ると合一間隔低減を示し、一方、減少集団に関してはこの逆も言える。

【0178】

米国およびタイにおける流行は、指数関数的に増大しつつある。米国およびタイにおけるサブタイプB流行に関する合一データは、流行病データと一致する。タイにおけるサブタイプB流行に関する合一データは、流行病データから予測されるよりも早い。この矛盾を説明し得る理由として考えられるものとしては、例えばHIV-1の多重導入の存在（この点に関する系統発生的証拠はない）、約7年間のタイにおけるHIV-1検出の非存在、ならびにHIV-1サブタイプEおよびBにお

ける env 遺伝子に関する突然変異率の差が挙げられる。

【0179】

再構築の単位

再構築のこの単位は、再構築される先祖ウイルス配列（即ち、状態）状態に関する。再構築の3つの考え得る単位が存在する：即ち、ヌクレオチド、アミノ酸またはコドンである。一実施態様では、個々のヌクレオチドの状態が再構築され、次にこの再構築を基礎にしてアミノ酸配列が確定される。別の実施態様では、アミノ酸先祖状態が直接再構築される。典型的実施態様では、進化のコドンモデルを用いる見込みベースの手法を用いて、コドンが再構築される。進化のコドンモデルは、コドンの頻度を、そして暗示的にあるヌクレオチドを別のヌクレオチドに置換する確率を考慮に入れる。言い換えれば、それは、単一モデルにおいてヌクレオチドおよびアミノ酸置換の両方を組入れる。当業者に理解されるように、これを実行し得るコンピュータプログラムが利用可能であり、あるいは容易に開発され得る。

【0180】

先祖状態を概算するための限界または合同見込みの使用

限界または合同見込みを用いて、先祖状態を概算し得る。限界および合同見込みは、系統発生樹における他の節点での先祖状態の概算方法に基づいて、異なる。任意の特定の系統樹に関しては、根の配列アラインメント上の所定部位の先祖状態が、例えばAである確率は、異なる方法で確定し得る。

【0181】

ヌクレオチドがアデニン (A) であるという見込みは、より高い父子（即ち先祖ウイルス配列、始祖またはMRCAにより近い節点）がアデニンを有するか、シトシン (C) か、グアニン (G) かまたはチミン (T) を有するかにかかわらず、確定し得る。これは先祖状態がAであるという限界見込みである。

あるいは、ヌクレオチドがAであるという見込みは、前記の節点がAであるか、Cか、GかまたはTであるかによって確定し得る。この概算はAの合同見込みで、その部位に関する他のすべての先祖再構築を伴う。

【0182】

合同見込みは、全系統樹に沿ったすべての先祖状態が確定される必要がある場合には、好ましい方法である。ある所定節点で最も考えられる状態を確立するためには、限界見込みが好ましくは用いられる。特定部位での不確定の場合、先祖状態の見込み概算は、ある状態が統計学的に別のものより良好であるか否かを試験するのを可能にする。2つの考え得る先祖状態が統計学的に異なる見込みを有さない場合、あるいは一方が終わりとなる場合には、すべての考え得る配列を構築する多数の部位に亘る多重状態は望ましくない。しかしながらすべての組合せの見込みをコンピューター処理し、ランク付けし得て、そしてある判定基準地より上のものだけを用いる。例えば、A、C、G、Tに関する異なる見込みを各々有する配列上の2つの部位が考えられる。

【0183】

L (A) L (C) L (G) L (T) *

*Lは、 $-1nL$ (負の対数見込み) を表し、したがって小さいほど考えられ得る。

部位1 3 2 1.5 1

部位2 10 7 5 1

16の考え得る配列形状が存在し、小野のそれ自体の対数見込みを有する。即ち、核塩基に関する対数見込みの合計を単純にする：

【0184】

【表1】

AA13	CA12	GA11.5	TA11
AC10	CC9	GC8.5	TC8
AG8	CG7	GG6.5	TG6
AT4	CT3	GT2.5	TT2

見込み順に、ランキングを示す：

TT、GT、CT、AT、TG、CG、CG、AG、TC、GC、CC、AC、TA、GA、CA、AA

【0185】

最初の4つの配列は、第二部位にTを有する。これは、大範囲に亘って拡散されるその部位での見込みに起因し、この部位にT以外の任意のヌクレオチドを有する確率を非常に低くする。しかしながら部位1では、任意のヌクレオチドが全く同様の見込みで与えられる傾向がある。この種のランキングは、変異が考慮される場合に捜し出すための考え得る配列の数を減じる一方法である。

【0186】

再構築先祖状態における前記の変異は、進化過程の確率的性質のために、ならびに典型的に用いられるその過程の蓋然的モデルのために、生じる変異を取り扱う。変異の別の供給源は、配列のサンプリングに起因する。サンプリングが先祖状態再構築に如何に影響するかを試験する一方法は、既存データ組におけるジャックナイフ再サンプリングを実施することである。これは、配列の何らかの部分（例えば半分）の取替えを伴わずに無作為に欠失し、そして先祖状態を再構築することを包含する。あるいは、先祖状態を一組のブートストラップ樹の各々に関して概算し、そして特定ヌクレオチドが概算された回数を所定部位に関する先祖状態として記録し得る。ブートストラップ樹はブートストラップデータを用いて生成するが、しかし先祖状態再構築は、元のデータでのブートストラップ樹を用いる。

【0187】

進化の異なるモデルを用いて、根底節点に関する先祖状態を再構築し得る。モデルの例は既知であり、多数のレベルに関して選定し得る。例えば進化のモデルは、いくつかの発見的手段により、または先祖配列に関する最高見込み（すべての部位に亘る見込みを合計することにより得られる）を生じるものを抜き取ることにより選定し得る。あるいは先祖状態は、進化のすべてのモデル全体の各部位で、合計獲得見込みのすべてで、そして最大見込みを有する選定先祖状態で、再構築される。

【0188】

実施例3.

HIV-1サブタイプC CTLアミノ酸コンセンサスエピトープの保存を分析した。エ

ピトープの総数は395であった。以下の表は、CサブタイプCTLコンセンサス配列との各循環性ウイルス配列の類似性の結果を要約する。HIV-1サブタイプCenvタンパク質（配列番号4）に関する確定先祖ウイルス配列は、最高得点（98.48%）を有する。切頭化配列を用いたために、いくつかの系統に関する得点が65%より低いことに留意。

【0189】

【表2】

配列名	全アミノ酸	コンセンサスに対するCTLの%
cCanc95-mod1	389	98.48%
cBR.92BR025	376	95.19%
cBI.BU910717	363	91.90%
cIN.21068	368	93.16%
cIN.301905	370	93.67%
cMW959.U08453	358	90.63%
cBW.96BW1210	365	92.41%
cBI.BU910316	367	92.91%
cZAM176.U86778	352	89.11%
cMW965.U08455	364	92.15%
cZAM174.16.U86768	351	88.86%
c84ZR085.U88822	322	81.52%
cSN.SE364A	370	93.67%
cMW960.U08454	365	92.41%
cBI.BU910812	368	93.16%
cET.ETH2220	358	90.63%
cBI.BU910518	361	91.39%
cIN.94IN11246	361	91.39%
cBW.96BW15B03	359	90.89%
cDJ.DJ259A	355	89.87%
cBI.BU910213	365	92.41%

【0190】

【表3】

配列名	全アミノ酸	コンセンサスに対するCTLの%
cBW.96BW01B03	362	91.65%
cIND760.L07655	255	64.56%
cIN.301904	372	94.18%
cSO.SM145A	354	89.62%
cCHN19.AF268277	356	90.13%
cIND747.L07653	255	64.56%
cBW.96BW0402	364	92.15%
cBI.BU910611	367	92.91%
cBI.BU910423	359	90.89%
cBW.96BW17B05	355	89.87%
cBW.96BW0502	367	92.91%
cUG.UG268A2	372	94.18%
cZAM18.L22954	365	92.41%
cIN.301999	368	93.16%
c91BR15.U39238	371	93.92%
cDJ.DJ373A	361	91.39%
cBI.BU910112	369	93.42%
c93IN101.AB023804	365	92.41%
cBW.96BW16B01	361	91.39%
cBW.96BW11B01	361	91.39%
cINDiananc66	363	91.90%

【0191】

実施例4.

マッカク属のサルにおいて増殖したサル免疫不全ウイルスに関して、先祖配列再構築を実施した。マッカク属サルを相対的に同種のSIV接種物に感染させ、チ

チャレンジさせた。感染後3年までにウイルス配列を得て、最大見込み系統発生分析を用いてMRCAを推定するために用いた。その結果生じた配列を、接種物のコンセンサス配列と比較した。MRCA配列は、ウイルス接種物と97.4%同一であることが判明した。この数値は、5つのグリコシル化部位での収束が除去されると、98.2%に改善した。この収束は、動物における増殖のための組織培養彼のウイルスの再適応のためであった (Edmonson et al., J. Virol. 72:405-14(1998))。MRCA配列およびコンセンサス配列は、ヌクレオチドレベルで1.5%異なることが判明した。図3は、サル免疫不全ウイルスMRCA系統発生の確定を示す。

【0192】

実施例5.

HIV-1サブタイプB先祖ウイルスenv遺伝子配列の生物学的活性を試験するための実験を実施した。HIV-1サブタイプB先祖ウイルスenv遺伝子配列をコードする核酸配列を、長い(160~200塩基)オリゴヌクレオチドから収集した。収集遺伝子をANC1と名づけた。ANC1 HIV-1-B Envの生物学的活性を、補助受容体結合およびシンシチウム形成検定で評価した。確定および化学合成HIV-1サブタイプB先祖gp160Env配列を保有するプラスミドpANC1、またはHIV-1サブタイプB89.6gp160Envを含有する陽性対照プラスミドを、COS7細胞中にトランスフェクトした。これらの細胞は、高効率で外来DNAを取込み、発現し得るので、したがって他の細胞に対する提示のためのウイルスタンパク質を産生するためにルーチンに用いられる。

【0193】

次にトランスフェクト化COS7細胞を、2つの主要HIV-1補助受容体タンパク質CCR5またはCXCR4のうちの1つを発現するGHOST細胞と混合した。CCR5は、感染の初期にHIVにより用いられる優性受容体である。CXCR4は、感染の後期に用いられ、後者受容体の使用は疾患の発症と一次的に関連する。COS7-GHOST-補助受容体+細胞を次に、巨細胞形成に関しては光学顕微鏡により、そしてウイルスEnvタンパク質の発現に関してはHIV-Env-特異的抗体染色および蛍光検出によりモニタリングした。

【0194】

ANC1 Envを発現する細胞は、HIV-特異的抗体との結合、蛍光検出に基づいて発現されることが、そしてCCR5補助受容体の存在下で多核化巨細胞の形成を引き起こすが、CXCR4補助受容体の場合は起こさないことが示された。陽性対照89.6EnvはCCR5およびCXCR4の両方を用い、いずれかの補助受容体を発現する細胞を有するシンシチウムを形成した。したがってANC1 Envタンパク質は、補助受容体結合およびシンシチウム形成により生物学的に活性であることが示された。

【0195】

実施例6.

コンセンサス配列が配列中の各部位での最も一般的ヌクレオチドまたはアミノ酸残基の配列を表すため、最大見込み系統発生再構築は伝統的コンセンサス配列決定とは異なる。したがってコンセンサス配列は、偏りのあるサンプリングを受ける。特にコンセンサス配列の確定は、多数の試料が同一配列を有する場合には偏向される。さらにコンセンサス配列は実在するウイルス配列である。

これに対比して、最大見込み系統発生分析は、各位置での各ヌクレオチドの頻度のみをベースにした最新の共通先祖の配列を確定しないため、偏向試料の影響をあまり受けないと考えられる。確定先祖ウイルス配列は、サンプリングした循環性ウイルスの共通の先祖であるウイルスである実際のウイルスの概算である。

【0196】

先祖配列を確定するための方法のうち最も簡単なものでは、配列アラインメント上の単一に関して、節点間の変化の総数が最小にされるように、ヌクレオチドが先祖節点に割り当てられる。このアプローチは「最も最節約型の再構築」と呼ばれる。最大見込みの原理に基づいた代替的方法是、系統発生の場合に観察された配列を獲得する確率を最大にするように、節点にヌクレオチドを割り当てる。系統発生は、ヌクレオチド置換の確率を特定する進化のモデルを用いることにより構築される。最大見込み系統発生は、観察されたデータを生じる最高確率を有するものである。

【0197】

図5を参照すると、先祖ウイルス配列（例えば始祖配列または最新の共通先祖配列（MRCA））の確定の最節約的方法および最大見込み方法の比較が示されてい

る。最も節約的再構築（「MP」）は、先祖分枝点（即ち節点）で不明瞭な状態を生じるという望ましくない問題を有し得る。この例では、この節点からの2つの子孫配列は、配列中の特定の位置にアデニン（A）またはグアニン（G）を有する。この部位での先祖配列に関する最も節約的再構築（「MP再構築」）は、この位置でAまたはG（「R」と記号を付す）であり得るため、不明瞭である。これに對比して、最大見込み系統発生分析は、配列進化についての知識を適用する。例えば見込み分析は、部分的に、他の変異体中の同一位置でのヌクレオチドの同定に頼る。したがってこの例では、隣接節点の変異体も同一位置にGを有するため、GからAへの突然変異は、AからGへの突然変異より起こりそうである。

【0198】

図6を参照すると、別の例は、最新共通先祖を確定するためのこれらの方法における差を示す。この例では、7ヌクレオチドの12の配列が示されている。これらの配列は、示された進化歴を共有する。これらの配列から算定したコンセンサス配列は、CATACTGである。パネルAでは、確定先祖節点の最大見込み再構築は、GATCCTGとして示されている。その他の確定配列は、他の内節点に隣接して示されている。パネルBでは、同一節点での最も節約的再構築が示されている。図示されているように、最も節約的再構築は、コンセンサス配列GAWCCTGを予測する。

【0199】

この場合、「W」は、AまたはTが第三位置に存在することが等しく可能であることを示す符号である。同様に、他の最も節約的再構築は、種々の内節点に示されている。第7内節点では、最後にヌクレオチドは符号「V」で示されており、A、CまたはGが存在し得ることをあらわす。この例では、コンセンサス配列は少なくとも2つの部位（第1および第4位置）で、MRCAに関して最大見込み—または最節約—確定配列と異なることにも留意されたい。

【0200】

前記から、説明のために本発明の特定の実施態様を本明細書中に記載してきたが、本発明の精神および範囲を逸脱しない限り種々の修正がなされ得る、と理解される。本明細書中に引用した出版物および特許出願はすべて、個々の出版物ま

たは特許出願の各々が参照により組入れられることを特定の且つ個別に示された場合のように、それらの記載内容は、参照により本明細書中に含まれる。前記の発明は理解を平易にするために図面および実施例により多少詳細に記載してきたが、しかし、本発明の教示にかんがみて、添付の特許請求の精神または範囲を逸脱しない限りある種の変更および修正がそれになされ得ることは、当業者には容易に明らかになる。

【0201】

【表4】

表4 (配列番号: 1)

1	ATGCGCGTGA	AGGGCATCCG	CAAGAACTAC	CAGCACCTGT	GGCGCTGGGG
51	CACCATGCTG	CTGGGGATGC	TGATGATCTG	CTCCGGGGCC	GAGAAGTGTG
101	GGGTGACCGT	GTAATACGGC	GTGCCCGTGT	GGAAGGAGGC	CACCACCACC
151	CTGTTCTGCC	CCAGCGACGC	CAAGGCTTAC	GACACCGAGG	TCCACACCGT
201	GTGGGCCACC	CACGCTCGCG	TGCCACCCGA	CCCCAACCCC	CAGGAGGTGG
251	TGCTGGAGAA	CGTGACCGAG	AACCTCAACA	TGTGGAAGAA	CAACATGGTG
301	GAGCAGATGC	ACGAGGACAT	CATCAGCCTG	TGGGACCAGA	GCCTGAAGCC
351	CTGCGTGAAG	TTAACCCCCC	TGTGCGTGAC	CCTGAACTGC	ACCGACGACC
401	TGCGCACCAA	CGCCACCAAC	ACCACCAACA	GCAGCGCCAC	CACCAACACC
451	ACCAGCAGCG	GCGGCGGCAC	GATGGAGGGC	GAGAAGGGCG	AGATCAAGAA
501	CTGCAGCTTC	AACGTGACCA	CCAGCATCCG	CGACAAGATG	CAGAAGGAGT
551	ACGCCCTGTT	CTACAAGCTG	GACGTGGTGC	CCATCGACAA	CGACACAAAC
601	AACACCAACA	ACAACACCAG	CTACCGCCTC	ATCAACTGCA	ACACCAAGCGT
651	GATCACCACG	GCCTGCCCCA	AGGTGAGCTT	CGAGCCCATC	CCCATCCACT
701	ACTGCACCCC	CGCCGGCTTC	GCCATCCTGA	AGTGCAACGA	CAAGAAGTTC
751	AACGGCACCG	GCCCCGTGAC	CAACGTGAGC	ACCGTGACGT	GCACCCACGG
801	CATCCGCCCC	GTGGTGAGCA	CCCAGCTGCT	GCTGAACGCC	AGCCTGGCCG
851	AGGAGGAGGT	GGTGATCCGC	AGCGAGAACT	TCACCGACAA	CGCCAAGACC
901	ATCATCGTGC	AGCTGAACGA	GAGCGTGGAG	ATCAACTGCA	CGCGTCCCAA
951	CAACAACACC	CGCAAGAGCA	TCCCCATCGG	CCCTGGCCGC	GCCCTGTACG
1001	CCACCGGCAA	GATCATCGGC	GACATCCGCC	AGGCCCACTG	CAACCTGTCC
1051	CGAGCCAAAGT	GGAACAACAC	CCTGAAGCAG	ATCGTGACCA	AGCTGCCCGA
1101	GCAGTTCGGC	AACAACAAGA	CCACCATCGT	GTTCAACCAAG	AGCAGCGGCG
1151	GGGACCCCGA	GATCGTGATG	CACAGCTTCA	ACTGCGGGCG	CGAATTCTTC
1201	TACTGCAACA	GCACCCAGCT	GTTCAACAGC	ACCTGGCACT	TCAACGGCAC
1251	CTGGGGCAAC	AACAACACCG	AGCGCAGCAA	CAACGCGGCC	GACGACACCG
1301	ACACCATCAC	CCTGCCCTGC	CGCATCAAGC	AGATCATCAA	CATGTGGCAG
1351	GAGGTGGGCA	AGGCCATGTA	CGCCCCCCCC	ATCAGCGGCC	AGATCCGCTG
1401	CAGCAGCAAC	ATCACCAGGC	TGCTGCTGAC	TCGAGACGGC	GGCAACAACG
1451	AGAACACCAA	CAACACCGAC	ACCGAGATCT	TCCGCCCCCG	GGGCGGCGAC
1501	ATGCGCGACA	ACTGGGCGAG	CGAGCTGTAC	AAGTACAAGG	TGGTGAAGAT
1551	CGAGCCCCCTG	GGCGTGGGCC	CCACCAAGGC	CAAGCGCCGC	GTGGTGACGC
1601	GCGAGAAGCG	CGCCGTGGGC	ATGCTGGGCG	CCATGTTCCCT	GGGCTTCCTG
1651	GGCGCCGCGG	GCAGCACCAT	GGCGCCGCGC	AGCATGACCC	TGACCGTGCA
1701	GGCCCGCCAG	CTGCTGAGCG	GCATCGTGCA	GCAGCAGAAC	AACCTGCTGC
1751	GCGCCATCGA	GGCCAGCAG	CACCTGCTGC	AGCTGACCGT	GTGGGGCATC
1801	AAGCAGCTGC	AGGCCCGCGT	GCTGGCCGTC	GAGCGGTACC	TGAAGGACCA
1851	GCAGCTGCTG	GGCATCTGGG	GCTGCAGCGG	CAAGCTGATC	TGCACACCGC
1901	CGGTGCCCTG	CAACGCCAGC	TGGAGCAACA	AGAGCCTGGA	CAAGATCTGG
1951	AACAACATGA	CCTGGATGGA	GTGGGAGCGC	GAGATCGACA	ACTACACCGG
2001	CCTGATCTAC	ACCCTGATCG	AGGAGAGCCA	GAACCAAGCAG	GAGAAGAACG
2051	AGCAGGAGCT	GCTGGAGCTG	GACAAGTGGG	CCAGCCTGTG	GAATGGTTC
2101	GATATCACCA	ACTGGCTGTG	GTACATCAAG	ATCTTCATCA	TGATCGTGGG
2151	CGGCCTGGTG	GGCCTGCGCA	TGCTGTTCCG	CGTGTGAGC	ATCGTGAACC
2201	GCGTGGCGCA	GGGCTACAGC	CCCCTGAGCT	TCCAGACCCG	CCTGCCCGCC
2251	CCCCGCGGCC	CCGACCGCCC	CGAGGGCATC	GAGGAGGAGG	GCGGCGAGCG
2301	CGACCGCGAC	CGCAGCGGSC	GCCTGGTGAA	CGGCTTCCTG	GCCCTGATCT
2351	GGGACGACCT	GCGCAGCCTG	TGCCTGTTCA	GCTACCACCG	CCTGCCCGAC
2401	CTGCTGCTGA	TGCTGGCCCG	CATCGTGAGG	CTGCTGGGCC	GGCGCGGCTG
2451	GGAGGCCCTG	AAGTATTGGT	GGAACCTGCT	GCAGTACTGG	AGCCAGGAGC
2501	TGAAGAACAG	CGCCGTGAGC	CTGCTGAACG	CCACCGCCAT	CGCCGTGGCC
2551	GAGGBCACCG	ACCGCGTGAT	CGAGGTGGTG	CAGCGGCCCT	GCCGCGCCAT
2601	CCTGCACATC	CCCCGCGCGA	TCCGCCAGGG	CCTGGAGCGC	GCCCTGCTGT
2651	GA				

【0202】

【表5】

表5 (配列番号: 2)

MRVKGIRKNY	QHLWRWGTM	LGMLMICSAA	EKLWVTVYYG	VPVWKEATTT	LFCASDAKAY
DTEVHNWVAT	HACVPTDFNP	QEVVLENVTE	NFNMWKNNMV	EQMHEDIISL	WDQSLKPCVK
LTPLCVTLNC	TDDLRTNATN	TTNSSATTNT	TSSGGGTMEG	EKGEIKNCSF	NVTTSIRDKM
QKEYALFYKL	DVVPIIDNDN	NTNMNTSYRL	INCNTSVITQ	ACPKVSFEPI	PIHYCTPAGF
AILKCNDDKF	NGTGPCTNVS	TVQCTHGIRP	VVSTQLLLNG	SLAEDEVVIR	SENFTDNAKT
IIVQLNESVE	INCTRPNNNT	RKSIPIGPGR	ALYATGKIIG	DIRQAHCNLS	RAKWNNTLKQ
IVTKLREQFG	NNKTTIVFNQ	SSGGDPEIVM	HSFNCGGEFF	YCNSTQLFNS	TWHFNGTWGN
NNTERSNNAA	DDNDTITLPC	RIKQIINMWQ	EVGKANYAPP	ISGQIRCSSN	ITGLLLTRDG
GNNENTNNTD	TEIFRPGGGD	MRDNWRSELY	KYKVVKIEPL	GVAPTKAKRR	VVQREKRAVG
MLGAMFLGFL	GAAGSTMGAA	SMTLTVQARQ	LLSGIVQQQN	NLLRAIEAQQ	HLLQLTVWGI
KQLQARVLAV	ERYLKDQQLL	GIWGCSGXLI	CTTAVPWNAS	WSNKSLDKIW	NNMTWMEWER
EIDNYTGILIY	TLIEESQNQQ	EKNEQELLEL	DKWASLWNNF	DITNWLWYIK	IFIMIVGGVL
GLRIVFAVLS	IVNRVRQGY	PLSFQTRLFA	PRGPDRPEGI	EEEGGERDRD	RSGLVNGFL
ALIWDLLRSL	CLFSYHRLRD	LLLIVARIVE	LLGRRGWEAL	KYWWNLLQYW	SQELKNSAVS
LLNATAIAVA	EGTDRVIEWV	QRACRAILHI	PRRIRQGLER	ALL	

【0203】

【表6】

表6 (配列番号: 3)

ATGCGGGTGATGGGCATCCTGCGGAACGCCAGCAGTGGTGGATCTGGGGCATCCTGGGC
 TTCTGGATGCTGATGATCTGCAGCGTGATGGGCAACCTGTGGGTGACCGTGACTACGGC
 GTGCCCCGTGTGGAAGGAGGCCAAGACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTAC
 GAGCGGAGGTGCACAACGTGTGGGCCACCCACGCCCTGCGTGCCCCACCGACCCCAACCCC
 CAGGAGATGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACCACATGGTG
 GACCAGATGCACGAGGACATCATCAGCCTGTGGGACCGAGCCTGAAGCCCTGCGTGGAAG
 CTGACCCCCCTGTGCGTGACCCCTGAACCTGCACCAACCGTGACCAACACCAACAACAAC
 AACACCAGCATGGGCGGCGAGATCAAGAAGTGCAGCTTCAACATCACCACCGAGCTGCGG
 GACAAGAAGCAGAAGGTGTACGCCCTGTCTACCGGCTGGACATCGTGCCCCCTGAACGAG
 AACAGCAACAGCAACAGCAGCGAGTACCGGCTGATCAACTGCAACACCAGCGCCATCACC
 CAGGCCCTGCCCCAAGGTGAGCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGC
 TACGCCATCCTGAAGTGCAACAACAAGACCTTCAACGGCACCGGCCCTGCAACAACGTG
 AGCACCGTGCACTGCACCCACGGCATCAAGCCCGTGGTGAGCACCCAGCTGCTGCTGAAC
 GGCAGCCTGGCCGAGGAGAGATCATCATCCGAGCGGAGAACCTGACCAACAACGCCAAG
 ACCATCATCGTGCACTGAACGAGAGCGTGAGATCGTGTGCACCCGGCCCAACAACAAC
 ACCCGGAAGAGCATCCGGATCGGCCCGCGGACAGACCTTCTACGCCACCGGCGACATCATC
 GCGGACATCCGGCAGGCCCACTGCAACATCAGCGAGAAGGAGTGGAAACAAGACCCCTGCAG
 CGGGTGGGCAAGAAGCTGAAGGAGCACTTCCCCAACAAGACCATCAAGTTCCGAGCCCAGC
 AGCGGCGGCGACCTGGAGATCACCACCCACAGCTTCAACTGCCGGGCGAGTTCTTCTAC
 TGCAACACCAGCAAGCTGTTCAACAGCACCTACAACAGCACCAACAACGGCACCCACAGC
 AACAGCACCATCACCTGCCCCTGCCGGATCAAGCAGATCATCAACATGTGGCAGGGCGTG
 GGCCGGGCCATGTACCGGGCCCCCATCGCCGGCAACATCACCTGCAAGAGCAACATCACC
 GGCTGCTGCTGACCCGGCAAGCGGCAACCAACAACACCACCGAGACCTTCCGGCCCC
 GCGGCGGCGGACATGCGGGACAACCTGGCGGAGCGAGCTGTACAAGTACAAGGTGGTGGAG
 ATCAAGCCCCCTGGGCGTGGCCCCCACCAGGGCCAAGCGGCGGGTGGTGGAGCGGAGAAG
 CGGGCCGTGGGCATCGGGCCCGTGTTCCTGGGCTTCCTGGGCGCCGCGGCGACCATG
 GGCGCCGCCAGCATCACCTGACCGTGACGGCCCGGCGAGCTGCTGAGCGGCATCGTGACG
 CAGCAGAGCAACCTGCTGCGGGCCATCGAGGCCCGAGCAGCAGCATGCTGCAGCTGACCGTG
 TGGGGCATCAAGCAGCTGCAGACCCGGGTGCTGGCCATCGAGCGGTACCTGAAGGACCAG
 CAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGG
 AACAGCAGCTGGAGCAACAAGAGCCAGGACGACATCTGGGACAACATGACCTGGATGACG
 TGGGACCGGGAGATCAGCAACTACACCGACACCATCTACCGGCTGCTGGAGGACAGCCAG
 AACCAGCAGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACAGCTGGAAGAACCTGTGG
 AACTGGTTCGACATCACCACCTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGC
 GGCCTGATCGGCCTGCGGATCATCTTCGCCGTGCTGAGCATCGTGAACCGGGTGCGGCAG
 GGCTACAGCCCCCTGAGCTTCCAGACCCCTGACCCCCAACCCCCGGGGCCCCGACCGGCTG
 GGCGGCATCGAGGAGGAGGGCGCGGAGCAGGACCGGGACCGGAGCATCCGGCTGGTGAGC
 GGCTTCCTGGCCCTGGCCTGGGACGACCTGCGGAGCCTGTGCCTGTTAGCTACCACCGG
 CTGCGGGACTTCATCCTGATCGCCGCCCCGGGGCGTGAACCTGCTGGGCGGAGCAGCCTG
 CGGGGCTGCAAGCGGGGCTGGGAGGCCCTGAAGTACCTGGGCAGCCTGGTGCAGTACTGG
 GGCCTGAGCTGAAGAAGAGCGCCATCAGCCTGCTGGACACCATCGCCATCGCCGTGGCC
 GAGGGCACCGACCGGATCATCGAGCTGGTGACGCGATCTGCGGGCCATCCGGAACATC
 CCCCCGCGATCCGSCAGGGCTTCGAGGCCGCCCTGCAGTGA

[0204]

[表7]

表7 (配列番号: 4)

MRVMGILRNCQQWWIWGILGFWMMLMICSVMGNLWVTVYYGVFWKEAKTT
 LFCASDAKAYEREVHNVWATHACVPTDPNPQEMVLENVTENFNMWKNDMV
 DQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVNTNNNNNTSMGGEIKN
 CSFNITTELKDKKQKVYALFYRLDIVPLNENSNSNSSEYRLJNCNTSAIT
 QACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGK
 PVVSTQLLNGSLAEEEEIIIRSENLTNNAKTIIVHLNESVEIVCTRPNN
 TRKSIRIGPGQTFYATGDIIGDIRQAHCNISEKEWNKTLQRVGKKLKEHF
 PNKTIKFEPSSGGDLIITHSFNCRGEFFYCNTSKLFNSTYNSTNNGTTS
 NSTITLPCRKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGN
 TNNITETFRPGGDMRDNRSELYKVKVEIKPLGVAPTEAKRRVVEREK
 RAVGIGAVFLGFLGAAGSTMCAASITLTVQARQLLSGIVQQSNLLRAIE
 AQOHMLQLTWVGIKQLQTRVLAIERYLKDQQLLGIWGCSGKLICTTAVPW
 NSSWSNKSQDDIWDNMTWMQWDREISNYTDTIYRLLEDSONQOEKNEKDL
 LALDSWKNLWNNFDITNWLWYIKIFIMIVGGLIGLRIFIIVLSIVNRVRQ
 GYSPLSFQTLTPNPRGPDRLLGGIEEGGEQDRDRSIRLVSGFLALAWDDL
 RSLCLFSYHRLRDFILIAARGVNLLGRSSLRGLQRGWEALKYLGSLVQYW
 GLELKSAISLLDTIAIAVAEGTDRIIELVQRICRAIRNIPRRIRQGF
 ALQ

【0205】

【表8】

表8 (配列番号: 5)

ATGAGAGTGAAGGGGATCAGGAAGAACTATCAGCACTTGTGGAGATGGGG
 CACCATGCTCCTTGGGATGTTGATGATCTGTAGCGCCGCCGAGAGCTGT
 GGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCACCAACC
 CTGTTCTGCGCCAGCGAGCCCAAGGCTTACGACACCGAGGTCCACAACGT
 GTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCAAGGAGTGG
 TGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTG
 GAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCC
 CTGCGTGAAGTTAACCCCCCTGTGCGTGACCCCTGAACCTGCACCGACGACC
 TGCGCACCAACGCCACCAACACCAACACAGCAGCGCCACCAACACACCC
 ACCCAGCAGCGCGCGGCGGACGATGGAGGGCGAGAAGGCGGAGATCAAGAA
 CTGCAGCTTCAACGTGACCACCAAGCATCCGCGACAAGATGCAGAGGAGT
 ACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACAC
 AACACCAACAAACACCAAGCTACCGCCTCATCAACTGCAACACCAAGCGT
 GATCACCCAGGCGCTGCCCCAAGGTGAGCTTCGAGCCCATCCCATCCACT
 ACTGCAACCCCTGCGCGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTC
 AACCGCACCGGCGCCCTGCAACCAAGTGAGCAGCAGTGCAGTGCACCCACGG
 CATCCGCCCCGTGGTGAGCAGCCAGCTGTGCTGAACGGCAGCCTGGCGG
 AGGAGGAGGTGGTGATCCGCGAGCGAGAATTACCGCAACAGCCCAAGACC
 ATCATCGTGAGCTGAACGAGAGCGTGGAGATCAACTGCAGCGCTCCCAA
 CAACAAACACCCGCAAGAGCATCCCATCGGCCCTGGCGCGCCCTGTACG
 CCACCGGCAAGATCATCGCGACATCCGCCAGGCCCACTGCAACCTGTGCG
 CGAGCCAAAGTGGAAACAACACCTGAAGCAGATCGTGACCAAGCTGCGCGA
 GCAGTTCGGCAACAACAAGACCAACCATCGTGTTCACCAAGAGCAGCGCGG
 GCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGCGCGCAATTCTTC
 TACTGCAACAGCACCAGCTGTTCAACAGCAGCTGGGCACTTCAACGGCAC
 CTGGGGCAACAACAACACCGAGCGCAGCAACAACGCGCGCGACGACAACG
 ACACCATCACCCCTGCCCCGCGATCAAGCAGATCATCAACATGTGGCAG
 GAGGTGGGCAAGGCCATGTACGCCCCCCCCATCAGCGGCCAGATCCGCTG
 CAGCAGCAACATCACCGGCGCTGCTGCTGACTCGAGACGGCGGCAACAACG
 AGAACACCAACAACACCGACACCGAGATCTTCGCCCCGGGGCGCGGAC
 ATCGCGCAAACTGCGCGAGCGAGCTGTACAAGTACAGGTGGTGAAGAT
 CGAGCCCTGGGCGTAGCACCCACCAAGGCAAGAGAGAGTGGTGCAGA
 GAGAAAAGCGCAGTGGGAATGCTAGGAGCTATGTTCTTGGGTCTTG
 GGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTGACCGTACA
 GGCCAGACAATTATGCTGCTGATAGTGACGAGCAGCAACAATCTGCTGA
 GGGCTATTGAGGCGCAACAGCATCTCTTGCAACTCAGCTCTGGGCGATC
 AAGCAGCTCCAGGCAAGAGTCTTGCTGTGGAAAGATACCTAAAGGATCA
 GCAGCTCCTGGGGATTGTTGGGTTGCTCTGGAAAACCTCATCTGCACCACTG
 CTGTGCTTGGGAATGCTAGCTGGAGCAACAAGAGCCTGGACAAGATCTGG
 AACCAATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCGG
 CCTGATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACG
 AGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
 GATATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGG
 CGGCTGCTGGGCTGCGCATCGTGTTCGCGCTGCTGAGCATCGTGAACC
 GCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCACCTGCCAGCC
 CCGAGGGGACCCGACAGGCCCGAAGGAATCGAAGAAGAGGTGGAGAGAG
 AGACAGAGACAGATCCGCTCGATTAGTGAATGGATTCTTAGCACTTATCT
 GGGACGACCTGCGSAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGCGAC
 TTACTCTTGATTGTAGCGAGGATTGTGGAACCTCTGGGACGCGAGGGGTG
 GGAGGCCCTCAAAATATTGGTGGAACTCTCTGAGTACTGGAGTCAAGAAC
 TAAAGAATAGCGCGGTGAGCCTGCTGAACGCCACCGCCATCGCGGTGGCC
 GAGGGCACCGACCGCGTATCGAGGTGGTGCAGCGCGCTGCGCGCCAT
 CCTGCACATCCCCCGCGCATCCGCCAGGGCCTGGAGCGCGCCCTGCTGT
 GA

【0206】

【表9】

表9 (配列番号: 6)

ATGAGAGTGATGGGGATACTGAGSAATTGTCAACAATGGTGGATATGGGG
 CATCCTAGGCTTTTGGATGCTAATGATTTGTGACGTGATGGGCAACCTGT
 GGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACC
 CTGTTCTGCGCCAGCGACGCCAAGGCCCTACGAGCGGGAGGTGCACAACTG
 GTGGGCCACCCACGCTTGGCTGCCACCGACCCCAACCCCCAGGAGATGG
 TGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTG
 GACCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCC
 CTGCGTGAAGCTGACCCCTGTGCTGACCTGAACTGCACCAACGTGA
 CCAACACCAACAACAACAACAACACGAGCATGGGCGGCGAGATCAAGAAC
 TGCAGCTTCAACATCACCAACGAGCTGCGGGACAAGAAGCAGAAGGTGTA
 CGCCCTGTCTTACCGGCTGGACATCGTGCCCTGAACGAGAACAGCAACA
 GCAACAGCAGCGAGTACCGGCTGATCAACTGCAACACGAGCGCCATCACC
 CAGGCTGCCCAAGGTGAGCTTGCACCCATCCCATCCACTACTGCGC
 CCCCBCCGGCTACGCCATCTGAAGTGCAACAACAAGACCTTCAACCGCA
 CCGGCCCTGCACAACGTGAGCAGCGTGAGTGACCCACGGCATCAAG
 CCGTGTGTGAGCACCAGCTGCTGCTGAACGCGAGCCTGGCCGAGGAGGA
 GATCATCATCCGAGCGAGAACCTGACCAACAACGCCAAGACCATCATCG
 TGCACCTGAACGAGAGCGTGGAGATCGTGTGCACCCGGCCCAACAACAAC
 ACCCGGAAGAGCATCCGGATCGGCCCGCGGACAGCCTTCTACGCCACCGG
 CGACATCATCGGCGACATCCGGCAGGCCCATGCAACATCAGCGAGAAGG
 AGTGAACAAGACCCCTGCAGCGGTGGGCAAGAAGCTGAAGGAGCACTTC
 CCAACAAGACCATCAAGTTCGAGCCAGCAGCGCGCGGACCTGGAGAT
 CACCACCCACAGCTTCAACTGCCGGGGCGAGTTCTTCTACTGCAACACCA
 GCAAGCTGTTCAACAGCACCTACAACAGCACCAACAACGGCACCACCAGC
 AACAGCACCATCACCCCTGCCCTGCCGATCAAGCAGATCATCAACATGTG
 GCAGGGCGTGGGCGGGCCATGTACGCCCGCCCGCATCGCCGGCAACATCA
 CCTGCAAGAGCAACATCACCGGCGCTGCTGCTGACCCGGGACGGCGCAAC
 ACCAACAACAACACCGAGACCTTCCGGCCCGGGCGGGCGGACATGCGGGA
 CAACTGGCGGAGCGAGCTGTACAAGTACAAGGTGGTGGAGATCAAGCCCC
 TGGGCGTAGCACCCACTGAGGCAAAAAGGAGAGTGGTGGAGAGAGAAAA
 AGAGCAGTGGGAATAGGAGCTGTGTTCTTGGGTTCTTGGGAGCAGCAGG
 AAGCACTATGGGCGCGGCGTCAATAACGCTGACGGTACAGGCCAGACAAT
 TATTGTCTGGTATAGTGCAACAGCAAGCAATTTGCTGAGGGCTATAGAG
 GCGCAACAGCATATGTGCAACTCACGGTCTGGGGCATTAAGCAGCTCCA
 GACAAGAGTCTCTGGCTATAGAAAGATACCTAAGGATCAGCAGCTCCTGG
 GCATTTGGGGCTGCTCTGGAAACTCATCTGCACCACTGCTGTGCTTGG
 AACTCTAGCTGGAGCAACAAGAGCCAGSACGACATCTGGGACAACATGAC
 CTGGATGCAGTGGGACCGGGAGATCAGCAACTACCCGACACCATCTACC
 GGCTGCTGGAGGACAGCCAGAACCAGCAGGAGAAGAACGAGAAGGACCTG
 CTGGCCCTGGACAGCTGGAAGAACCTGTGGAACCTGGTTGACATCACCAA
 CTGGCTGTGTTACATCAAGATCTTCATCATGATCGTGGCGGCTGTATCG
 GCCTGCGGATCATCTTCGCCGTGCTGAGCATCGTGAACCGGGTGCAGGAG
 GGCTACAGCCCCCTGAGCTTCCAGACCTTACCCCAACCCGAGGSGACC
 CGACAGGCTCGGAGGAATCGAAGAAGAAGGTGGAGAGCAAGACAGAGACA
 GATCCATTGATTAGTGAGCGGATTCTTAGCACTGGCCTGGGACGACCTG
 CGGAGCCTGTGCTCTTCAGCTACCAACCGATGAGAGACTTCATATTGAT
 TGCAGCCAGAGGGTGGGAACCTCTGGGACGAGCAGTCTCAGGGGACTGC
 AGAGGGGGTGGGAAGCCCTTAAGTATCTGGGAAGTCTTGTGCACTATTGG
 GGCTTGGAGCTAAAAAGAGTGCTATTAGCCTGCTGGACACCATCGCCAT
 CGCCCTGCGCGAGGCGACCGGATCATCGAGCTGGTGCAGCGGATCT
 GCGGGCCATCCGGAACATCCCCCGCGGATCCGGCAGGGCTTCGAGGCG
 GCGCTGCAGTGA

【図面の簡単な説明】

【図1】

図1は、HIV-1の系統発生的分類を示す。丸で取り囲まれた節点は、HIV-1主要

群（群M）および主要群クレードA～G、J、AGIおよびAGの先祖状態に近い。

【図2】

図2は、HIV-1サブタイプBの系統発生関係および系統樹上の確定されたサブタイプB先祖節点の配置を示す。HIV-1サブタイプDの系統発生的関係は、外集団として示されている。

【図3】

図3は、アカゲザルへの感染後3年までにSIV接種物に関して最大見込み再構成を用いて最新の共通の先祖の先祖ウイルス配列再構成を示す。コンセンサス配列および最も近年の共通先祖配列はヌクレオチド配列で1.5%異なることが判明した。

【図4】

図4は、先祖ウイルス配列を用いたデジタルワクチンの開発の一例を提供する。

【図5】

図5は、「大半の貧弱な再構成」法と「最大見込み再構成法」の比較を示す。

【図6】

図6は、「大半の貧弱な再構成」法と「最大見込み再構成法」の別の比較を示す。

【図7】

図7は、pJW4304SV40/EBVベクターのマッピングを示す。

【図8】

図8は、HIV-1サブタイプCの系統発生関係および系統樹上の確定されたサブタイプC先祖節点の配置を示す。

【配列表】

SEQUENCE LISTING

<110> UNIVERSITY OF WASHINGTON

<120> AIDS ANCESTRAL VIRUSES AND VACCINES

<130> 16336-13-1PC

<140> PCT/US01/

<141> 2001-02-16

<150> USSN 60/183,659

<151> 2000-02-18

<160> 23

<170> PatentIn Ver. 2.1

<210> 1

<211> 2651

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ancestral
HIV-1 group M, subtype B, env sequence.

<400> 1

```

atgcgcgtga agggcatccg caagaactac cagcaoctgt ggcgctgggg caccatgctg 60
ctgggggatgc tgatgatctg ctccgcggcc gagaagcgtg gggtagccgt gtactacggc 120
gtgcccgtgt ggaaggaggc caccaccacc ctgttctgcg ccagcgacgc caaggcttac 180
gacaccgagg tccacaacgt gtgggccacc cagccctgcg tgcccaccga cccaacccc 240
caggaggtgg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300
gagcagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag 360
ttaaccccc tgtgcgtgac cctgaactgc accgacgacc tgcgcaccaa cgcacccaac 420
accaccaaca gcagcgccac caccaacacc accagcagcg gcggcgggac gatggagggc 480
gagaaggggc agatcaagaa ctgcagcttc aacgtgacca ccagcatccg cgacaagatg 540
cagaaggagt acgcccctgt ctacaagctg gacgtggtgc ccatcgacaa cgacaacaac 600
aacaccaaca acaacaccag ctaccgcctc atcaactgca acaccagcgt gatcaccacc 660
gcctgcccc aggtgagctt cgagcccctc cccatccact actgcacccc cgcgggcttc 720
gccatcctga agtgcaacga caagaagttc aacggcaccc gccctgcac caacgtgagc 780
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 840
agcctggcgg aggaggaggt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 900
atcatcgtgc agctgaacga gagcgtggag atcaactgca cgcgtcccaa caacaacacc 960
cgcaagagca tcccctcgg ccctggcgcg gccctgtacg ccaccggcaa gatcatcggc 1020
gacatccgct aggcccactg caacctgtcg cgagccaagt ggaacaacac cctgaagcag 1080
atcgtgacca agctgcgcga gcagttcggc aacaacaaga ccaccatcgt gttcaaccag 1140
agcagcggcg gcgaccccga gatcgtgatg cacagcttca actgcggcgg cgaattcttc 1200

```

```

tactgcaaca gcacccagct gttcaacagc acctggcact tcaacggcac ctggggcaac 1260
aacaacacccg agcgcagcaa caacggccgc gacgacaacg acaccatcac cctgcccctgc 1320
cgcatacaagc agatcatcaa catgtggcag gaggtgggca aggccatgta cggccccccc 1380
atcagcggcc agatccgctg cagcagcaac atcaccggcc tgcctgctgac tcgagacggc 1440
ggcaacaacg agaacaccaa caacaccgac accgagatct tccgccccgg gggcgcgac 1500
atgcgcgaca actggcgagc cgagctgtac aagtacaagg tggatgaagat cgagccccctg 1560
ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgcctggggc 1620
atgctggggc ccattgttctt gggtttctct ggcgccggcg gcagcaccat gggcgccgc 1680
agcatgaccc tgaccgtgca ggcggccag ctgctgagcg gcacgtgca gcagcagaac 1740
aacctgctgc ggcacatcga ggcggcgag cactgtctgc agctgacct gtggggcatc 1800
aagcagctgc aggcggcggt gctggcgctg gagcggtacc tgaaggacca gcagctgctg 1860
ggcatctggg gctgcagcgg caagctgac tgcaaccacg cgggtgccctg gaacggcagc 1920
tggagcaaca agagcctgga caagatctgg aacaacatga cctggatgga gtgggagcgc 1980
gagatcgaca actacacggc cctgatctac accctgatcg aggagagcca gaaccagcag 2040
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggctc 2100
gatataccca actggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 2160
ggcctgcgca tcgtgttcgc cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 2220
ccctgagct tcacagcccg cctgcccgc ccccgcgcc ccgaccgccc cgagggcac 2280
gaggaggagg gcggcgagcg cgaccgagc cgcagcgggc gcctggtgaa cggcttctc 2340
gcctggtctt gggacgacct gcgcagcctg tgctgttca gctaccacg cctgcgcgac 2400
ctgctgctga tcgtggcccg catcgtggag ctgctgggccc ggcgcggtct ggaggccctg 2460
aagtattggt ggaacctgct gcagtactgg agccaggagc tgaagaacag cgcctgagc 2520
ctgctgaacg ccaccggcat cgcctggccc gagggcaccg acccgctgat cgagggtggt 2580
cagcgcgctt gcgcggcat cctgcacat ccccgccga tccgcccagg cctggagcgc 2640
gcctgctgt ga
2652

```

<210> 2

<211> 883

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ancestral
HIV-1 group M, subtype B, env sequence.

<400> 2

```

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
  1                      5                      10                     15

Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Lys
          20                      25                     30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
          35                      40                     45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
          50                      55                     60

```


His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	65	70	75	80
Gln	Glu	Val	Val	Leu	Glu	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys	85	90	95	
Asn	Asn	Met	Val	Glu	Gln	Met	His	Glu	Asp	Ile	Ile	Ser	Leu	Trp	Asp	100	105	110	
Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu	115	120	125	
Asn	Cys	Thr	Asp	Asp	Leu	Arg	Thr	Asn	Ala	Thr	Asn	Thr	Thr	Asn	Ser	130	135	140	
Ser	Ala	Thr	Thr	Asn	Thr	Thr	Ser	Ser	Gly	Gly	Gly	Thr	Met	Glu	Gly	145	150	155	160
Glu	Lys	Gly	Glu	Ile	Lys	Asn	Cys	Ser	Phe	Asn	Val	Thr	Thr	Ser	Ile	165	170	175	
Arg	Asp	Lys	Met	Gln	Lys	Glu	Tyr	Ala	Leu	Phe	Tyr	Lys	Leu	Asp	Val	180	185	190	
Val	Pro	Ile	Asp	Asn	Asp	Asn	Asn	Asn	Thr	Asn	Asn	Asn	Thr	Ser	Tyr	195	200	205	
Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	Val	Ile	Thr	Gln	Ala	Cys	Pro	Lys	210	215	220	
Val	Ser	Phe	Glu	Pro	Ile	Pro	Ile	His	Tyr	Cys	Thr	Pro	Ala	Gly	Phe	225	230	235	240
Ala	Ile	Leu	Lys	Cys	Asn	Asp	Lys	Lys	Phe	Asn	Gly	Thr	Gly	Pro	Cys	245	250	255	
Thr	Asn	Val	Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Arg	Pro	Val	Val	260	265	270	
Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Glu	Val	Val	275	280	285	
Ile	Arg	Ser	Glu	Asn	Phe	Thr	Asp	Asn	Ala	Lys	Thr	Ile	Ile	Val	Gln	290	295	300	
Leu	Asn	Glu	Ser	Val	Glu	Ile	Asn	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	305	310	315	320

Arg Lys Ser Ile Pro Ile Gly Pro Gly Arg Ala Leu Tyr Ala Thr Gly
 325 330 335
 Lys Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser Arg Ala
 340 345 350
 Lys Trp Asn Asn Thr Leu Lys Gln Ile Val Thr Lys Leu Arg Glu Gln
 355 360 365
 Phe Gly Asn Asn Lys Thr Thr Ile Val Phe Asn Gln Ser Ser Gly Gly
 370 375 380
 Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe
 385 390 395 400
 Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp His Phe Asn Gly
 405 410 415
 Thr Trp Gly Asn Asn Asn Thr Glu Arg Ser Asn Asn Ala Ala Asp Asp
 420 425 430
 Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
 435 440 445
 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln
 450 455 460
 Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
 465 470 475 480
 Gly Asn Asn Glu Asn Thr Asn Asn Thr Asp Thr Glu Ile Phe Arg Pro
 485 490 495
 Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
 500 505 510
 Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys
 515 520 525
 Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly Met Leu Gly Ala
 530 535 540
 Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 545 550 555 560
 Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val
 565 570 575

Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu
 580 585 590
 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu
 595 600 605
 Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 610 615 620
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser
 625 630 635 640
 Trp Ser Asn Lys Ser Leu Asp Lys Ile Trp Asn Asn Met Thr Trp Met
 645 650 655
 Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr Gly Leu Ile Tyr Thr Leu
 660 665 670
 Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu
 675 680 685
 Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn
 690 695 700
 Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val
 705 710 715 720
 Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg
 725 730 735
 Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Arg Leu Pro Ala Pro Arg
 740 745 750
 Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp
 755 760 765
 Arg Asp Arg Ser Gly Arg Leu Val Asn Gly Phe Leu Ala Leu Ile Trp
 770 775 780
 Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
 785 790 795 800
 Leu Leu Leu Ile Val Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly
 805 810 815
 Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr Trp Ser Gln
 820 825 830

Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn Ala Thr Ala Ile Ala
 835 840 845

Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val Val Gln Arg Ala Cys
 850 855 860

Arg Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg
 865 870 875 880

Ala Leu Leu

<210> 3

<211> 2561

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ancestral

HIV-1 group M, subtype C, env sequence.

<400> 3

atgcggtga tgggcaccc tgggaactgc cagcagtggt ggatctgggg catcctgggc 60
 ttctggatgc tgatgatctg cagcgtgatg ggcaacctgt gggtagacct gtactacggc 120
 gtgcccgtgt ggaaggaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac 180
 gagcgggagg tgcacaacgt gtggggccacc cagcctgctg tggccaccga ccccaacccc 240
 caggagatgg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa cgacatgggt 300
 gaccagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctggaag 360
 ctgaccccc tgtgctgac cctgaactgc accaactga ccaacaccaa caacaacaac 420
 aacaccagca tgggcggcga gatcaagaac tgcagcttca acatcaccac cgagctgctg 480
 gacaagaagg agaaggtgta cgccctgttc taccggctgg acatcgtgac cctgaacgag 540
 aacagcaaca gcaacagcag cgagtagcgg ctgatcaact gcaacaccag cgcacacacc 600
 caggcctgccc ccaaggtgag ctctgacccc atccccatcc actactgccc ccccgccggc 660
 tacgccatcc tgaagtgcac caacaagacc ttcaacggca ccggcccttg caacaacgtg 720
 agcaccgtgc agtgaccca cggcatcaag cccgtggtga gcacccagct gctgctgaac 780
 ggcagccttg ccgaggagga gatcatcatc cggagcgaga acctgaccaa caacgccaa 840
 accatcatcg tgcacctgaa cgagagcgtg gagatcgtgt gcacccggcc caacaacaac 900
 acccggaaga gcacccggat cggccccggc cagaccttct acgcccacgg cgacatcatc 960
 ggcgacatcc ggcaggccca ctgcaacatc agcgagaagg agtggaaaca gacctgcag 1020
 cgggtgggca agaagctgaa ggagcacttc cccaacaaga ccatcaagtt cgagccagc 1080
 agcggcgccg acctggagat caccacccac agcttcaact gccggggcga gttcttctac 1140
 tgcaacacca gcaagctgtt caacagcacc tacaacagca ccaacaacgg caccaccagc 1200
 aacagcacca tcacctgtcc ctgocggatc aagcagatca tcaacatgtg gcagggcgtg 1260
 ggccggggca tgtacgcccc ccccatcgcc ggcaacatca cctgcaagag caacatcacc 1320
 ggcctgctgc tgacccggga cggcgccaac accaacaaca ccaccgagac cttccggccc 1380
 ggccggcgcc acatgcggga caactggcgg agcgagctgt acaagtacaa ggtggtggag 1440

```

atcaagcccc tgggggtggc cccaccgag gccaaagcggc gggtggtgga gcgggagaag 1500
cggggcogtgg gcacggggcg cgtgttcctg ggcttctctg gcggcgccgg cagcaccatg 1560
ggcgccggcca gcacacccct gaccgtgcag gcccggcagc tgotgagcgg catcgtgcag 1620
cagcagagca acctgctgcg ggccatcgag gcccggcagc acatgctgca gctgaccgtg 1680
tggggcatca agcagctgca gacccgggtg ctggccatcg agcgggtacct gaaggaccag 1740
cagctgctgg gcacctgggg ctgcagcggc aagctgatct gcaccaccgc cgtgcccctg 1800
aacagcagct ggagcaacaa gagccaggac gacatctggg acaacatgac ctggatgcag 1860
tgggaccggg agatcagcaa ctacaccgac accatctacc ggctgctgga ggacagccag 1920
aaccagcagg agaagaacga gaaggacctg ctggccctgg acagctggaa gaacctgtgg 1980
aactggttcg acatcaccaa ctggctgtgg tacatcaaga tcttcacat gatcgtgggc 2040
ggcttgatcg gcctgcggat catcttcgcc gtgctgagca tcgtgaaccg ggtgcggcag 2100
ggctacagcc cctgagctt ccagacctg accccaacc cccggggccc cgaccggctg 2160
ggcgccatcg aggaggaggg cggcgagcag gacccggacc ggagcaccg gctggtgagc 2220
ggcttcctgg ccttggcctg ggacgacctg cggagcctgt gcctgttcag ctaccaccgg 2280
ctgcccgaact tcactctgat cggcgcccg ggctggaacc tgotggggcg gagcagcctg 2340
cggggcctgc agcggggctg ggaggccctg aagtacctg gcagcctggt gcagtactgg 2400
ggcttggagc tgaagaagag cgcctcago ctgctggaca ccacgccat cgcctgggcc 2460
gagggcaccg accggatcat cgagctggtg cagcggatct gccgggccc cgggaacatc 2520
ccccggcgga tccggcaggg ctteggaggc gccctgcagt ga 2562

```

<210> 4

<211> 852

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ancestral
HIV-1 group M, subtype C, env sequence.

<400> 4

```

Met Arg Val Met Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp
  1                      5                      10                     15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Ser Val Met Gly Asn
                20                      25                     30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
                35                      40                     45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Arg Glu Val
                50                      55                     60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
                65                      70                     75                     80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
                85                      90                     95

```

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asn Val Thr Asn Thr Asn Asn Asn Asn Thr Ser Met
 130 135 140
 Gly Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg
 145 150 155 160
 Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val
 165 170 175
 Pro Leu Asn Glu Asn Ser Asn Ser Asn Ser Ser Glu Tyr Arg Leu Ile
 180 185 190
 Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe
 195 200 205
 Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu
 210 215 220
 Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val
 225 230 235 240
 Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln
 245 250 255
 Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser
 260 265 270
 Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu
 275 280 285
 Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser
 290 295 300
 Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile
 305 310 315 320
 Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Glu Lys Glu Trp Asn
 325 330 335
 Lys Thr Leu Gln Arg Val Gly Lys Lys Leu Lys Glu His Phe Pro Asn
 340 345 350

Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr
 355 360 365
 Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser
 370 375 380
 Lys Leu Phe Asn Ser Thr Tyr Asn Ser Thr Asn Asn Gly Thr Thr Ser
 385 390 395 400
 Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
 405 410 415
 Trp Gln Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn
 420 425 430
 Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
 435 440 445
 Gly Asn Thr Asn Asn Thr Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp
 450 455 460
 Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu
 465 470 475 480
 Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val
 485 490 495
 Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe
 500 505 510
 Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr
 515 520 525
 Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn
 530 535 540
 Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val
 545 550 555 560
 Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr
 565 570 575
 Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu
 580 585 590
 Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser
 595 600 605

Gln Asp Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu
 610 615 620
 Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln
 625 630 635 640
 Asn Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp
 645 650 655
 Lys Asn Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile
 660 665 670
 Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile
 675 680 685
 Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro
 690 695 700
 Leu Ser Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Pro Asp Arg Leu
 705 710 715 720
 Gly Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile
 725 730 735
 Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser
 740 745 750
 Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala
 755 760 765
 Ala Arg Gly Val Asn Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu Gln
 770 775 780
 Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp
 785 790 795 800
 Gly Leu Glu Leu Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala
 805 810 815
 Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Leu Val Gln Arg
 820 825 830
 Ile Cys Arg Ala Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe
 835 840 845
 Glu Ala Ala Leu Gln
 850

<210> 5

<211> 2652

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Semi-optimized
ancestral viral sequences for HIV-1 subtypes B and
C.

<400> 5

```

atgagagtgga aggggatcag gaagaactat cagcacttgt ggagatgggg caccatgctc 60
cttgggatgtg tgatgatctg tagcgccgcc gagaagctgt gggtgaccgt gtactacggc 120
gtgcccgtgtg ggaaggaggc caccaccacc ctgttctgcg ccagcgacgc caaggcttac 180
gacaccgaggg tccacaacgt gtgggccacc cagccctgcg tgcccacoga ccccaacccc 240
caggaggtggg tgcctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300
gagcagatgc acgaggacat catcagcctg tgggaccaga gctgaagcc ctgctggaag 360
ttaaccccc tgctgctgac cctgaactgc accgacgacc tgccgaccaa cgccaccaac 420
accaccaaca gcagcgccac caccacacc accagcagcg gcggcggcac gatggagggg 480
gagaaggggcg agatcaagaa ctgcagcttc aacgtgacca ccagcatccg cgacaagatg 540
cagaaggagt acgcccgtgt ctacaagctg gacgtggtgc ccacgacaa cgacaacaac 600
aacaccaaca acaacaccag ctaccgcctc atcaactgca acaccagcgt gatcaccctg 660
gcctgcccc aagtgagctt cagagccatc cccatccact actgcacccc cgccgggttc 720
gccatcctga agtgcaacga caagaagttc aacggcaccg gccctgacac caactgagc 780
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 840
agcctggccg aggaggaggt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 900
atcatcgtgc agctgaacga gagcgtggag atcaactgca cgcgtcccaa caacaacacc 960
cgcaagagca tcccctcagg cctggccgcg gccctgtacg ccaccggcaa gatcatcgcc 1020
gacatccgca agcccactg caacctgtcg cggcccaagt ggaacaacac cctgaagcag 1080
atcgtgacca agctgcgga gcagttcggc aacaacaaga ccaccatcgt gttcaaccag 1140
agcagcggcg gcgaccccga gatcgtgatg cacagcttca actgcccggg cgaattcttc 1200
tactgcaaca gcacccagct gttcaacagc acctggcact tcaacggcac ctggggcaac 1260
aacaacaccg agcgcagcaa caacgcccgc gacgacaacg acaccatcac cctgccctgc 1320
cgcatcaagc agatcatcaa catgtggcag gaggtgggca aggccatgta cgccccccc 1380
atcagcggcc agatccgtg cagcagcaac atcaccggcc tgctgctgac tcgagacggc 1440
ggcaacaacg agaacacca caacaccgac accgagatct tccgccccgg gggcgcgac 1500
atgscgaca actggcgag cgagctgtac aagtacaagg tggggaagat cgagccctg 1560
ggcgtagcac ccaccaaggc aaagagaaga gtggtgcaga gaaaaaaag cgcagtggga 1620
atgctaggag ctatgttctt tgggttcttg ggagcagcag gaagcactat gggcgagcgc 1680
tcaatgacgc tgaccgtaca ggccagacaa ctattgtctg gtatagtga gcagcagaac 1740
aatctgctga gggctattga ggccgaacag catctgttgc aactcacagt ctggggcacc 1800
aagcagctcc aggcaagagt cctggctgtg gaaagatacc taaaggatca gcagctcctg 1860
gggatttggg gttgctctgg aaaactcacc tgcaccactg ctgtgocctg gaatgctagc 1920
tggagcaaca agagcctgga caagatctgg aacaacatga cctggatgga gtgggagcgc 1980
gagatcgaca actacaccgg cctgatctac accctgatcg aggagagcca gaaccagcag 2040
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactgggtc 2100

```

```

gatataccca actggctgtg gtacatcaag atcttcacatca tgatcgtggg cggcctgggtg 2160
ggcctgcgca tcgtgttcgc cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 2220
cccttgagct tccagaccca cctgocagcc ccgaggggac ccgacaggcc cgaaggaaac 2280
gaagaagaag gtggagagag agacagagac agatccggtc gattagtga tggattctta 2340
gcacttatct gggacgaact gcgagcctg tgcctcttca gctaccaccg cttgagcgac 2400
ttactcttga ttgtagcgag gattgtgga cttctgggac gcaggggggtg ggaggccctc 2460
aaatatttgt ggaatctcct gcagtactgg agtcaggaa taaagaatag cgcctgagc 2520
ctgctgaacg ccaccgccat cgcctggcc gagggcaccg acccgctgat cgaagggtg 2580
cagcgcgcct gccgcgccat cctgcacatc ccccgccgca tccgccaggg cctggagcgc 2640
gcctgtctgt ga                                     2652

```

<210> 6

<211> 2561

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Semi-optimized
ancestral viral sequences for HIV-1 subtypes B and
C.

<400> 6

```

atgagagtga tggggatact gaggaattgt caacaatggt ggatatgggg catcctaggg 60
ttttggatgc taatgatttg tgacgtgatg ggcaacctgt ggttgaccgt gtactacggc 120
gtgcccggtg ggaaggaggg caagaccacc ctgttctgag ccagcgacgc caaggcctac 180
gagcgggagg tgcacaacgt gtgggccacc cagcctgagc tgcccacgca ccccaacccc 240
caggagatgg tgctggagaa cgtgaccggg aacttcaaca tgtggaagaa cgacatgggt 300
gaccagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctgtaag 360
ctgacccccc tgtgcgtgac cctgaactgc accaacgtga ccaacaccaa caacaacaac 420
aacaccagca tgggcggcga gatcaagaac tgcagcttca acatcacca cagctgagg 480
gacaagaagc agaaggtgta cgcctgttgc taccggctgg acatcgtgac cctgaacgag 540
aacagcaaca gcaacagcag cgagtaccgg ctgatcaact gcaacaccag cgcctacacc 600
caggcctgac ccaaggtgag ctgcacccc atccccatc actactgccc ccccgccggc 660
tacgccatcc tgaagtgcac caacaagacc ttcaacggca ccggccctct caacaacgtg 720
agcaccgtgc agtgaccca cggcatcaag cccgtggtga gcacccagct gctgctgaac 780
ggcagcctgg ccgaggagga gatcatcatc cggagcgaga acctgaccaa caacgccaag 840
accatcatcg tgcacctgaa cgagagcgtg gagatcgtgt gcacccggcc caacaacaac 900
acccggaaga gcatccggat cggccccggc cagaccttct acgcccacgg cgacatcatc 960
ggcgacatcc ggcaggccca ctgcaacatc agcgagaagg agtggaaaca gacctgagc 1020
cgggtgggca agaagctgaa ggagcacttc cccaacaaga ccatcaagtt cgagccagc 1080
agcggcgggc acctggagat caccaccac agcttcaact gcggggcgga gttcttctac 1140
tgcaacacca gcaagctgtt caacagcacc tacaacagca ccaacaacgg caccaccagc 1200
aacagcacca tcacctgccc ctgcccggatc aagcagatca tcaacatgtg gcaggcggtg 1260
ggccgggcca tgtacgccc ccccatcgcc ggcaacatca cctgcaagag caacatcacc 1320
ggcctgctgc tgacccggga cggcggaac accaacaaca ccaccgagac cttccggccc 1380
ggcggcgggc acatgcggga caactggcgg agcgagctgt acaagtadaa ggtggtggag 1440
atcaagcccc tgggcgtagc acccactgag gcaaaaagga gagtgggtga gagagaaaaa 1500

```

```

agagcagtgg gaataggagc tgtgttccctt gggttcttgg gagcagcagg aagcactatg 1560
ggcgcggcgt caataacgct gacgggtacag gccagacaat tattgtotgg tatagtgcac 1620
cagcaaagca atttgctgag ggctatagag gcgcaacagc atatgttgca actcacgggc 1680
tggggcatta agcagctcca gacaagagtc ctggctatag aaagatacct aaaggatcag 1740
cagctcctgg gcatttgggg ctgctctgga aaactcatct gcaccactgc tgtgccttgg 1800
aactctagct ggagcaacaa gagccaggac gacatctggg acaacatgac ctggatgcag 1860
tgggaccggg agatcagcaa ctacaccgac accatctacc ggctgctgga ggacagccag 1920
aaccagcagg agaagaacga gaaggacctg ctggccctgg acagctggaa gaacctgtgg 1980
aactggttgc acatcaccaa ctggctgtgg tacatcaaga tcttcatcat gatcgtgggc 2040
ggcctgatcg gcctgcggat catcttcgcc gtgctgagca tcgtgaaccg ggtgcggcag 2100
ggctacagcc ccttgagctt ccagaccctt accccaaacc cgaggggacc cgacaggctc 2160
ggaggaatcg aagaagaagg tggagagcaa gacagagaca gatccattcg attagttagc 2220
ggattcttag cactggcctg ggacgacctg cggagccctg gcctcttcag ctaccaccca 2280
ttgagagact tcatattgat tgcagccaga gggtagggac ttctgggacg cagcagtctc 2340
aggggactgc agagggggtg ggaagccctt aagtatctgg gaagtcttgt gcagtattgg 2400
ggtctggagc taaaaaagag tgctattagc ctgctggaca ccacgccat cgcctgtggc 2460
gagggcaccg accggatcat cgagctggtg cagcggatct gcggggccat ccggaacatc 2520
ccccggcgga tccggcaggg cttcgaggcc gccctgcagt ga 2562

```

<210> 7

<211> 7

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
sequence-maximum likelihood reconstruction of
determined ancestral node.

<400> 7

gatcctg

7

<210> 8

<211> 7

<212> DNA

<213> Artificial Sequence

<220>

<221> variation

<222> (3)

<223> W can be an A or T

<220>

<223> Description of Artificial Sequence: Consensus
sequence, most parsimonious reconstruction of
determined ancestral node.

<400> 8
gawcctg

7

<210> 9
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Consensus
sequence, maximum likelihood reconstruction of
determined ancestral node.

<400> 9
gaacctg

7

<210> 10
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Consensus
sequence, maximum likelihood reconstruction of
determined ancestral node.

<400> 10
gaaactc

7

<210> 11
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Consensus
sequence, maximum likelihood reconstruction of
determined ancestral node.

<400> 11
gatactc

7

<210> 12

<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<221> variation
<222> (3)
<223> W can be an A or T

<220>
<223> Description of Artificial Sequence:Consensus
sequence, most parsimonious reconstruction of
determined ancestral node.

<400> 12
gawactc

7

<210> 13
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 13
catactc

7

<210> 14
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 14
catactt

7

<210> 15
<211> 7

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 15
catacta

7

<210> 16
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 16
catattg

7

<210> 17
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<221> variation
<222> {7}
<223> V can also be an A, C or G

<220>
<223> Description of Artificial Sequence:Consensus
sequence, most parsimonious reconstruction of
determined ancestral node.

<400> 17
catactv

7

<210> 18
<211> 7
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 18

catgctg

7

<210> 19

<211> 7

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 19

catactg

7

<210> 20

<211> 7

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 20

caagctg

7

<210> 21

<211> 7

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 21
catgctg

7

<210> 22
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

<400> 22
cttgctg

7

<210> 23
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence- maximum likelihood reconstruction of
determined ancestral node.

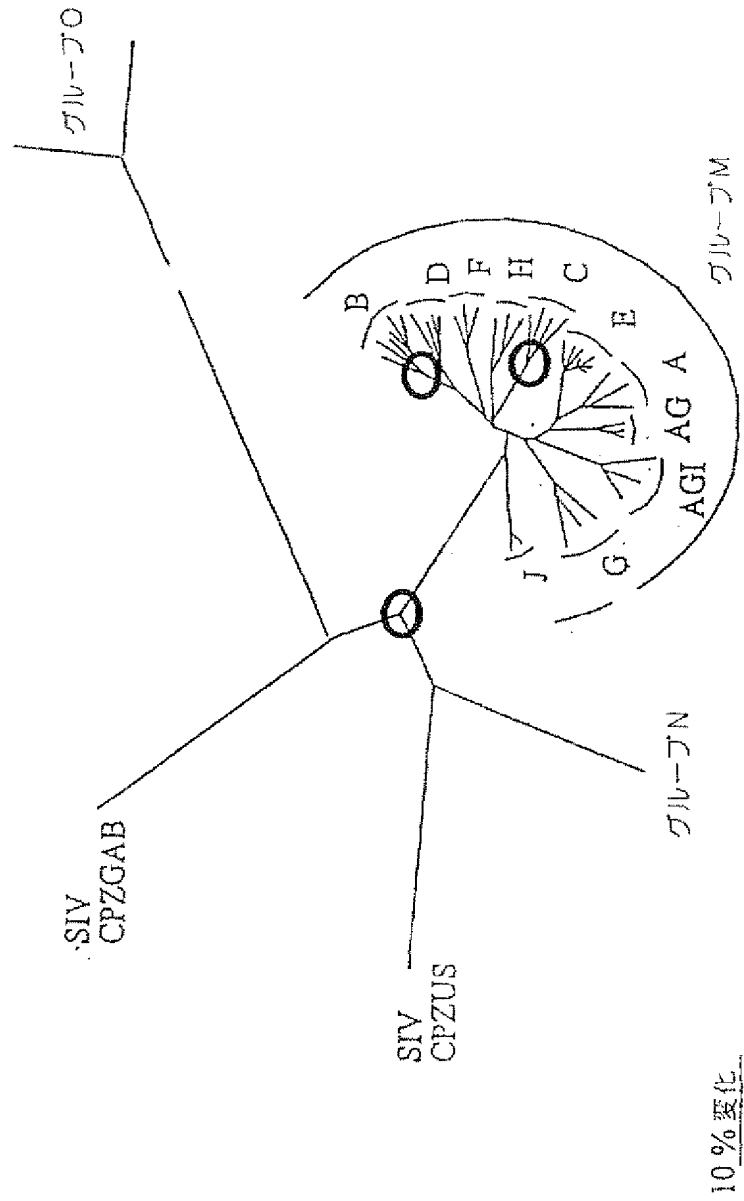
<400> 23
cttgctt

7

【図1】

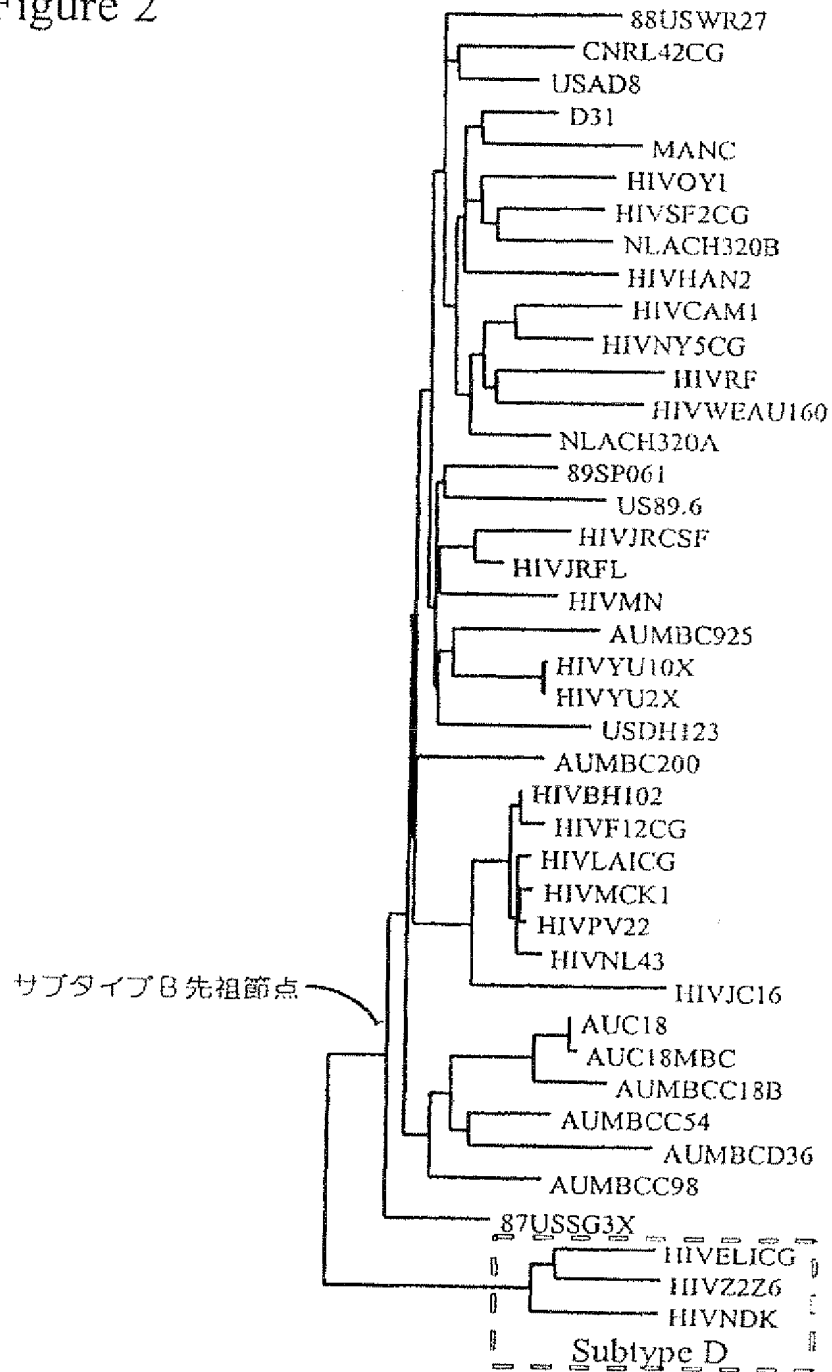
FIGURE 1

HIV-1の系統発生的分類



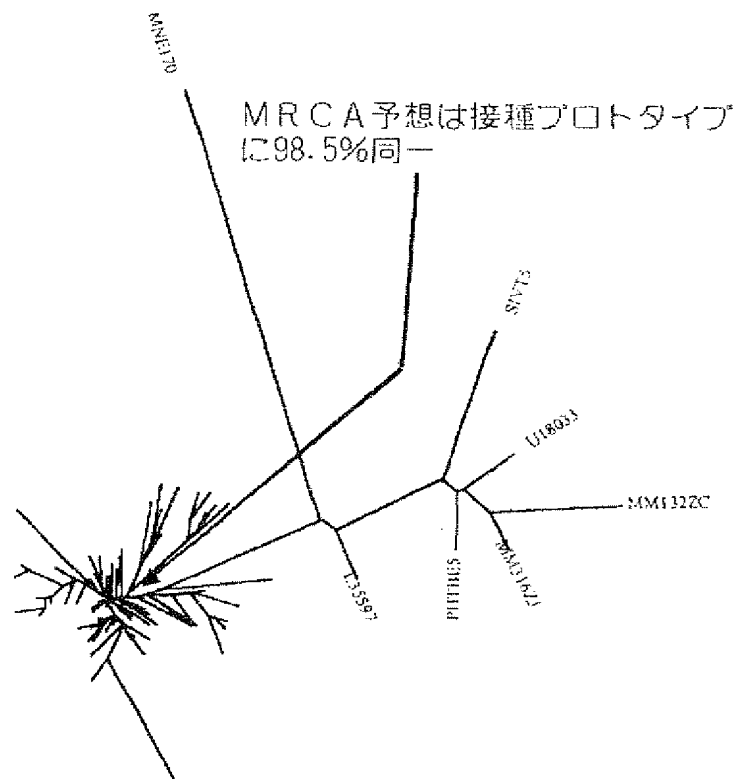
【図2】

Figure 2



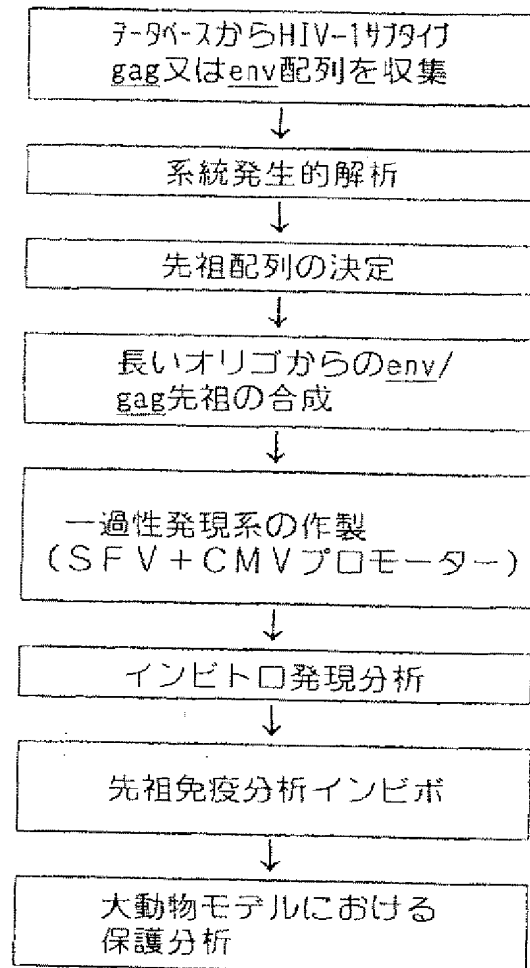
【図3】

Figure 3



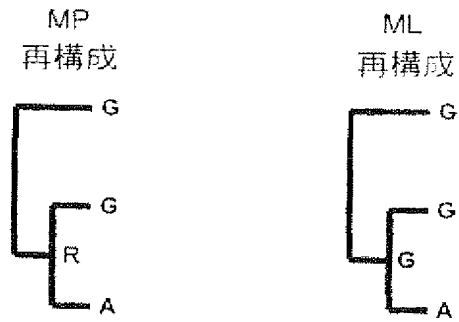
【図4】

Figure 4



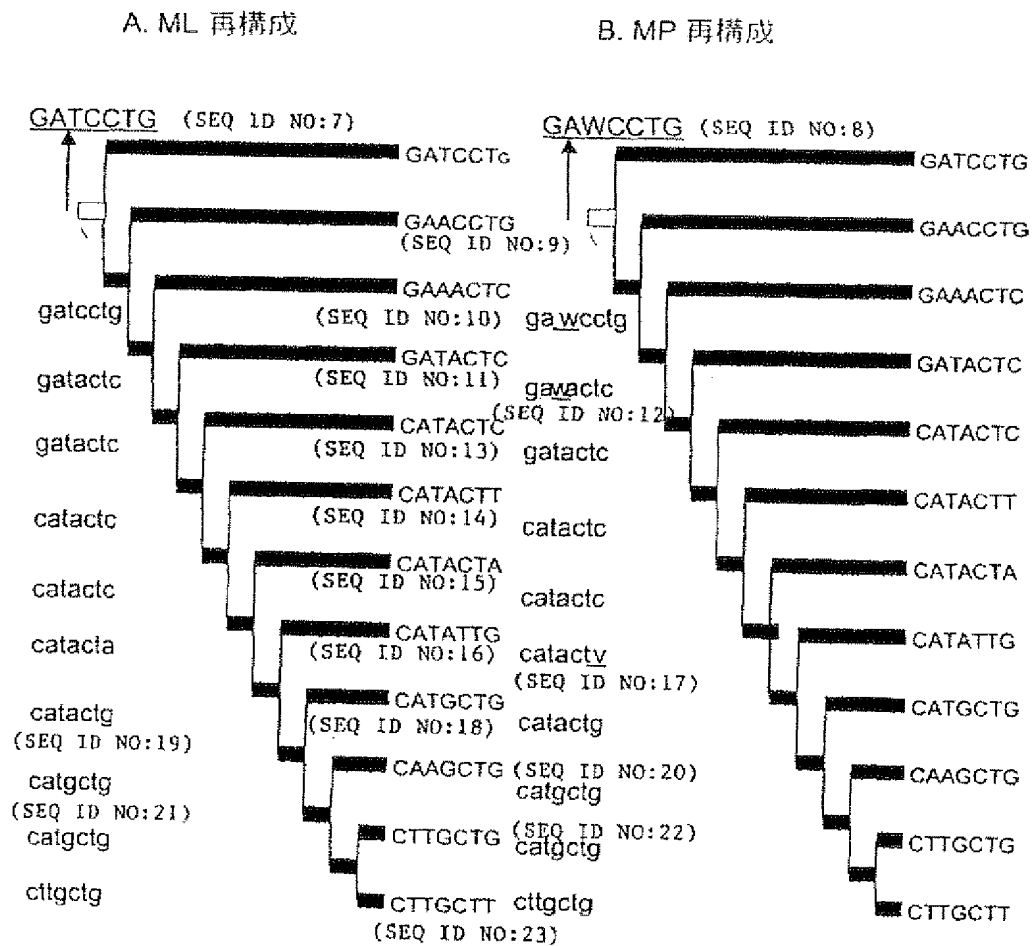
【図5】

Figure 5



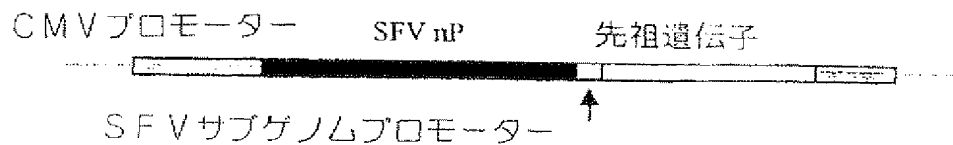
【図6】

Figure 6



【図7】

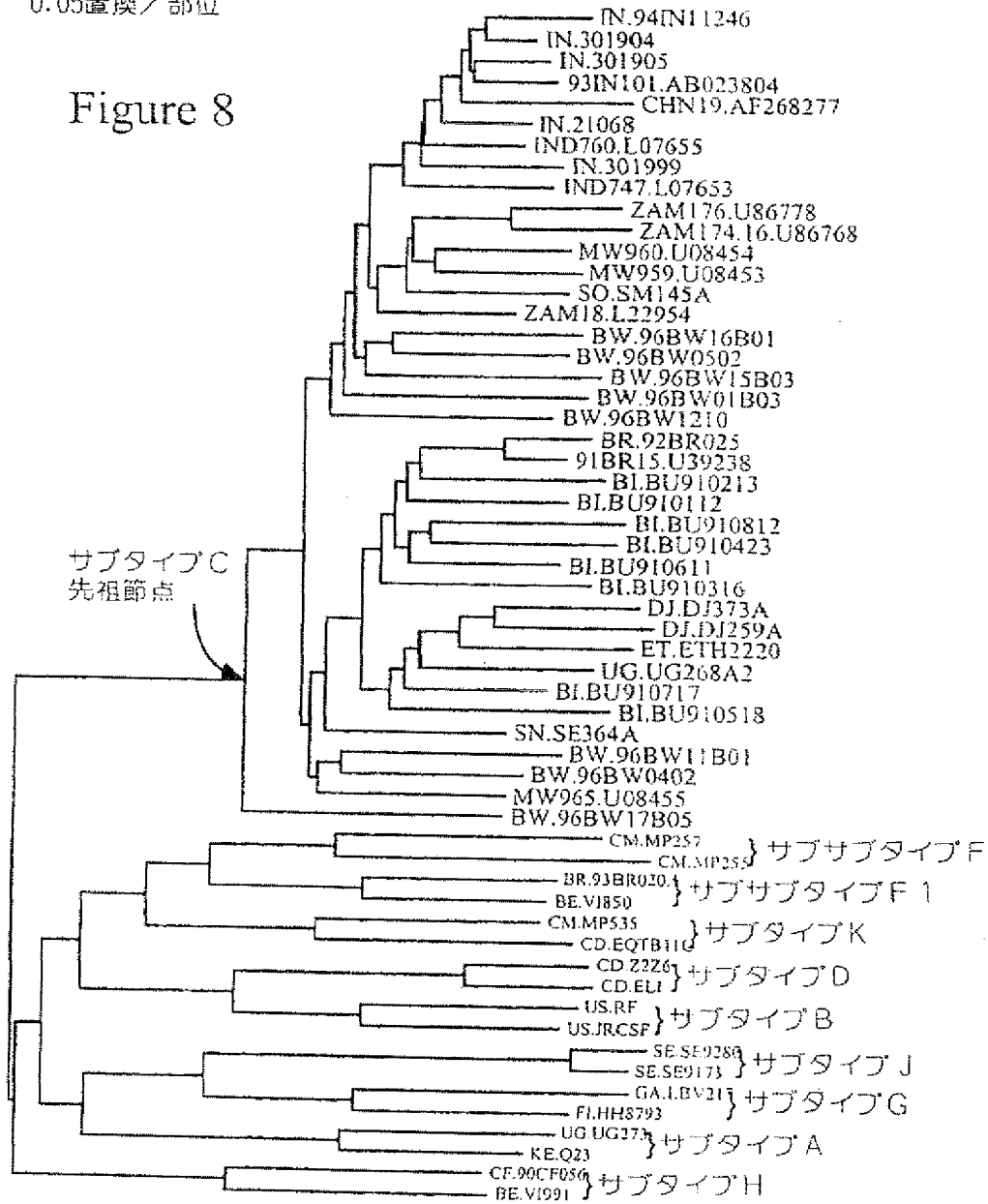
Figure 7



【図8】

0.05置換／部位

Figure 8



【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/05288

A. CLASSIFICATION OF SUBJECT MATTER

IPC(?) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
STIC database, Medline, WEST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y --- X	SCHADT et al. Computational advances in maximum likelihood methods for molecular phylogeny. Genome Research. March 1998, Vol. 8, No. 3, pages 222-233, see especially the "consensus" sequence.	7, 8, 10, 15, 16, 22, 23, and 31 1-6, 9, 11-14, 17-21, 24-30, and 32-43

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categorization of cited documents:	**	Inter document published after the international filing date or priority date and not in connection with the application but cited to understand the principle or theory underlying the invention
"A" Document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Z"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

30 JULY 2001

Date of mailing of the international search report

05 SEP 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-8580

Authorized officer

JEFFREY STUCKER

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/05288

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y --- X	ZHU et al. An African HIVI sequence from 1959 and implications for the origin of the epidemic. <i>Nature</i> . 05 February 1998, Vol. 391, No. 6667, pages 594-597, see entire document.	7, 8, 10, 15, 16, 22, 23, and 31 ----- 1-6, 9, 11-14, 17- 21, 24-30, and 32- 43
Y --- X	LOUWAGIE et al. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. <i>Journal of Virology</i> . January 1995, Vol. 69, No. 1, pages 263-271, see entire document.	7, 8, 10, 15, 16, 22, 23, and 31 ----- 1-6, 9, 11-14, 17- 21, 24-30, and 32- 43

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/05258

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 9.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/05288

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

A01N 57/18; A61K 36/00, 36/00, 36/21, 36/26, 36/42; C07H 21/02, 21/04; C07K 14/00, 16/00; C12N 5/00, 5/02, 5/06, 5/18, 15/09, 15/63; C12P 41/06; C12Q 1/70; G01N 33/53

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/5, 7.1, 69.1, 69.8, 325, 326, 349.1, 320.1; 514/2; 424/148.1, 160.1, 184.1, 186.1, 188.1, 206.1; 530/350, 388.35, 388.4; 588/28.1, 28.72

B. FIELDS SEARCHED

Minimum documentation searched
Classification System: U.S.

435/5, 7.1, 69.1, 69.8, 325, 326, 349.1, 320.1; 514/2; 424/148.1, 160.1, 184.1, 186.1, 188.1, 206.1; 530/350, 388.35, 388.4; 588/28.1, 28.72

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-5, drawn to a gene sequence.

Group II, claims 6-13 and 30-33, drawn to a protein.

Group III, claims 14-29, drawn to an expression construct.

Group IV, claims 34-35, drawn to an antibody.

Group V, claims 36-43, drawn to a method of preparing a sequence.

The inventions listed as Groups I-V do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The concept of tracing virus evolution to discern the relationships of different strains to and to generate the sequence of the ancestral strain(s) is known in the art. See the art cited in the search report. Therefore, the instant claims lack a special technical feature and unity of invention is lacking.

フロントページの続き

(51)Int.Cl. ⁷	識別記号	F I	キーワード (参考)
C 0 7 K	16/08	C 1 2 N	1/15
C 1 2 N	1/15		1/19
	1/19		1/21
	1/21	C 1 2 P	21/08
	5/10	C 1 2 N	15/00
// C 1 2 P	21/08		5/00
(81)指定国	EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, I T, LU, MC, NL, PT, SE, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, G M, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, B Z, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, J P, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, R O, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW		Z N A A A
(72)発明者	ラーン, ジェラルド エイチ, アメリカ合衆国, ワシントン 98346, キ ングストン, ノースイースト セカンド ストリート 11316		
(72)発明者	リ, フーシェン アメリカ合衆国, ワシントン 98115, シ アトル, # 3, ノースイースト セブンテ ィーフス ストリート 3818		

Fターム(参考) 4B024 AA01 BA35 BA51 BA61 CA04
CA07 DA01 DA02 DA05 DA06
DA11 EA02 EA04 GA01 GA11
HA08
4B064 AG26 AG27 AG31 AG32 CA02
CA05 CA10 CA11 CA12 CA19
CA20 CC24 DA03
4B065 AA01X AA26X AA57X AA80X
AA88X AA90X AA95Y AB01
AB04 BA01 BA08 CA24 CA25
CA44 CA45
4C085 AA03 BA65 EE01
4H045 AA10 AA11 AA20 AA30 BA09
BA41 CA05 DA75 DA76 DA86
EA31 FA72 FA74

